DOCUMENT RESUME

ED 227 508 CS 207 412

AUTHOR White, Edward M.
TITLE Comparison and Contrast: The 1977 California State University and Colleges Freshman English Equivalency Examination.
INSTITUTION California State Univ. and Colleges, Long Beach.; English Council of the California State Univ. and Colleges.
PUB DATE Oct 77
NOTE 91p.; For related documents, see ED 114 825-826 and CS 207 410-415.
PUB TYPE Reports - Descriptive (141) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS College Freshmen; Educational Testing; *Equivalency Tests; *Essay Tests; Higher Education; Objective Tests; Scoring; State Programs; Student Placement; Test Format; Testing Programs; *Test Interpretation; Test Norms; Test Reliability; *Writing Evaluation; *Writing Research
IDENTIFIERS California

ABSTRACT

The fifth administration of the California State University and Colleges Freshman English Equivalency Examination, a test consisting of an objective examination graded by the Educational Testing Service and an essay examination scored by California State University and College faculty, is analyzed in this report. After a short index of previous volumes and preface, the first section of the report contains a summary and fact sheet describing the test, its results, and studies being undertaken to examine the results. The next section recounts the actual administration of the test. The section on the examination's two essay questions consists of three subsections, two containing a report from the question leader, the question itself, the directions to graders, and six sample student papers for each of the questions, and the third listing the essay readers. Section four presents a statistical report on the examination scores, while section five presents the results of the third year of a 5-year follow-up study. The report's five appendixes cover a variety of topics, including budgets, the reliability of the examination, and an English placement test. (JL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED227508

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

COMPARISON AND CONTRAST

The 1977 California State University and Colleges

Freshman English Equivalency Examation

Edward M. White

Director of the Project
for
The California State University and Colleges

Project 77-329
Fund for Innovation and Improvement

October 1977

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Linda Bunnell Jones

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS207412

TABLE OF CONTENTS

Page

Short Index of Previous Volumes ii

Preface iii

I. Summary and Fact Sheet 1

II. Administration of the Test Program 6

III. The Essay Test 7

1. Question 1 8

a. Report by Question Leader: Dr. Eileen Lothamer 8
b. Question 1 8
c. Directions to Graders 8
d. Six Sample Student Papers 8

2. Question 2 14

a. Report by Question Leader: Dr. Rex Burbank 14
b. Question 2 14
c. Directions to Graders 14
d. Six Sample Student Papers 15

3. Table Chart/Reader Roster, 1977 19

IV. Statistical Report: Dr. Robert Bradley 20

V. Follow-up Studies 33

APPENDICES

A. Budget 48
An Exchange of Memos on the Budget 49

B. The 1977 English Equivalency Reliability Study, 1976-77 51

C. Son of EEE: The English Placement Test 57

D. Descriptive Statistics: Summary Reports 59

SHORT INDEX OF PREVIOUS VOLUMES
COMPARISON AND CONTRAST 1973, 1974, 1975, 1976

1. Summary and Fact Sheet
   - 1973: 1 – 13
   - 1974: 1 – 4
   - 1975: 1 – 4
   - 1976: 3 – 6

2. The Objective Test
   - 1973: 28 – 30
   - 1974: 13

3. The Essay Test: Reports, Questions, Sample Responses
   - 1973: 30 – 61
   - 1974: 15 – 39
   - 1975: 19 – 36
   - 1976: 11 – 26

4. Descriptive Statistics and Statistician Reports
   - 1973: 79 – 93, 132 – 167
   - 1974: VII – XX, xxiii – xxxvi
   - 1975: 51 – 68
   - 1976: 31 – 47, 84 – 97

5. Norm Samples
   - 1973 (7 CSUC campuses + UCLA and UCR): 51 – 61
   - 1974 (San Jose State, Stanford, UCR, UCSB, UCLA): 41-49, VII-XI
   - 1975 (San Francisco State; CSU, Long Beach): 54 – 62
   - 1976 (CSU, Northridge, CSU, Chico): 32 – 38

6. Follow-up Studies
   - 1975: 71 – 91
   - 1976: 48 – 63

7. Reports, Appendices and Charts
   - Equivalency Testing in College Freshman English
     - 1973: 94 – 125
   - Budgets
     - 1973: 126 – 131
     - 1974: XXI
     - 1975: 93
     - 1976: 64
   - Registration Form
     - 1974: 7
     - 1975: 7 – 10
   - English Council Committee Reports
     - 1974: 13
     - 1975: 12 – 16, 17 – 18
   - Organizational Charts
     - 1975: 94 – 98
     - 1976: 65
   - Media and Student Questionnaires
     - 1975: 99 – 100
   - Data Processing Procedures
     - 1975: 37 – 50
   - Four Issues for Faculty in Equivalency Testing
     - 1975: 101 – 111
   - The Working of a Controlled Essay Reading
     - 1976: 68 – 75

PREFACE

This report describes the fifth administration of The California State University and Colleges Freshman English Equivalency Examination. The development and initial implementation of the program are fully described in the first volume of *Comparison and Contrast,* which describes the 1973 program. The further development of the program is described in *Comparison and Contrast* 1974, 1975, and 1976. Since this 1977 volume will not repeat material already published in the first four reports, anyone seeking a full account of the development of the English Equivalency Examination should acquire the earlier four volumes from the Office of the Chancellor, The California State University and Colleges, 400 Golden Shore, Long Beach, California 90802.

Following the Table of Contents in this volume is a list of materials published in earlier reports that may be of particular interest to those receiving this report, which will go into detail only when new problems or new solutions occurred in 1977. This report does contain the essay questions that were used in 1977, the criteria for scoring, and six sample papers at the different score levels for each question. In addition, Appendix B contains some new material on the reliability that may be achieved by a controlled essay reading, information which may be of particular value to those institutions seeking to follow this model for obtaining reliable scoring of essays. Yet another section which may be of particular interest to those who have been following previous reports is a third-year report on the extensive follow-up studies of the 1973 test group. Finally, Appendix C speaks to the relation between English Equivalency and the new CSUC English Placement Test.

The English Equivalency Examination program has been made possible through the continued financial support of the Fund for Innovation and Improvement in the Instructional Process of The California State University and Colleges. This year for the fifth consecutive year all administrative, publication, and essay examination expenses have been covered by this fund. The continued support of the program by the California State University English Council, composed of English Department Chairs and freshmen English coordinators from all 19 campuses of the system, provides the program with a sound professional center.

Finally, it is worthy of note that, as the teaching and measurement of writing skills becomes more and more important at the college and university level, the English Equivalency Examination has served to involve a substantial number of university-level English faculty in the preparation and scoring of an essay examination according to the highest professional and psychometric standards. The continuing dialogue among the professionals involved in this program on all CSUC campuses has had a substantial effect on the teaching of writing and the teachers of writing throughout the state of California. It may turn out that this "in-service training" byproduct of English Equivalency will be the most important educational impact of the program. At the same time, we should not minimize the effect of the program upon the substantial numbers of students who have directly benefited from it.

LIST OF TABLES

Progress Report Number Three presents data concerning the progress of the original English Equivalency Examination participants as of the beginning of their fourth year of study. Tables are included which contain summarized and detailed information on students' general academic progress, accumulated units and grade-point average, intercampus transfer activity, and changes in major.

		Page
I.	Progress of Freshman Participants As of Fall 1976 by Sex	35
II.	Progress of Freshman Participants As of Fall 1976 by Ethnicity	36
III.	Mean Earned S.E.U. and GPA for Participants Enrolled Fall 1976	37
IV.	Mean Earned S.E.U. and GPA for June 1976 Graduates	38
V.	Mean Transferred and Total S.E.U. for Participants Enrolled Fall 1976 by Fall 1973 Enrolled Campus	39
VI.	Mean Transferred and Total S.E.U. for Participants Enrolled Fall 1976 by Fall 1973 Major	40
VII.	Mean Transferred and Total GPA for Participants Enrolled Fall 1976 by Fall 1973 Enrolled Campus	41
VIII.	Mean Transferred and Total GPA for Participants Enrolled Fall 1976 by Fall 1973 Major	42
IX.	Changes in Campus and Enrolled Major Fall 1973 to Fall 1976	43
X.	Intercampus Transfer Activity Fall 1973 to Fall 1976	44
XI.	Fall 1973 Enrolled Campus by Fall 1976 Enrolled Campus	45
XII.	Changes in Enrollment by Major Fall 1973 to Fall 1976	46
XIII.	Fall 1973 Major by Fall 1976 Major	47

SUMMARY AND FACT SHEET

The 1977 English Equivalency Examination proceeded smoothly, following, in general, the precedents established during the first four administrations. Enrollment continued to decrease. Speculation about the reasons for this decrease has led only to a determination that in 1978 redoubled efforts to publicize the benefits and possibilities of the examination for potential students will need to be made. Decreasing enrollment has, however, led to an increasing percentage of students passing the examination, so it does appear that the weakest students no longer attempt English Equivalency. This is as it should be.

Certainly, one major indication of the continued success of the program is the number of times it is mentioned as a model for new testing programs, both statewide and on individual campuses. With large-scale testing programs in English now underway in the CSUC, the English Equivalency model and experience has become an extremely important base upon which to develop new programs.

The following chart compares the number of applicants, and their success rates, across the five years of the English Equivalency Examination program. It shows the decline in number of enrollments, partially compensated for by the increasing percentage of passing scores.

STUDENTS APPLYING FOR ENGLISH EQUIVALENCY EXAMINATION CREDIT 1973-77

	Number Taking Test	Number Passing	Percent Passing	Semester Units Earned
1973	4,072	1,363	33.5	8,178
1974	3,639	1,036	28.5	6,216
1975	3,303	981	29.7	5,886
1976	2,866	1,093	38.1	6,558
1977	2,216	883	39.8	5,298
Five-Year Total	16,096	5,356	33.0	32,136

The multiple-choice test continued to be the CLEP Subject Examination, Analysis and Interpretation of Literature, developed and graded by the Educational Testing Service. The essay test was produced by two CSUC Test Development Committees and scored by The California State University and Colleges faculty at the controlled essay reading at Kellogg West on the campus of California State Polytechnic University, Pomona, June 11-15, 1977.

In addition to the test papers, readers also scored fifty-eight essay tests written by California State University, Long Beach freshmen, and eighty written by California State University, Fullerton freshmen. Dr. Robert Bradley, test officer at California State Polytechnic University, Pomona, was once again in charge of the statistical operation. As has been the case in the past, the controlled essay reading achieved an extremely high rate of consistency among the readers. A new procedure to record discrepant scores shows that of the 4,224 pairs of scores given by readers, only 302 were more than one point apart and required a third reading (149 for Question 1; 153 for Question 2). The essay test showed a significantly higher average score than last year's group achieved, no doubt in part because the ability level of the test group as a whole seems to have significantly increased. The essay readers in general felt that the 1977 essays showed enough of an improvement over the previous year's test to suggest that recent concern with high school composition instruction might be yielding encouraging results. The essay scoring followed closely the pattern of previous years, although standards were stiffened slightly. The four independent readings on a 6-point scale were added to yield a single essay test score on a scale from 0-24. Essay total scores of 14 or above when combined with passing scores on the machine-scored test allowed a student to pass.

For the first time since the English Equivalency Examination program began, a second level "placement" pass was established. Students passing the test at this level were exempted from the required CSUC English Placement Test, even though their scores were not high enough to deserve credit. Any student with a total essay score of 15 or higher received a placement pass, whatever the multiple-choice score; an essay score of 11 or higher earned this pass when combined with a multiple-choice score of 35 or higher. When we add the 1,104 students (49.8%) who received this placement, or exemption, pass to the 883 (39.8%) who received credit for freshman English through the English Equivalency Examination, approximately 90% of the 1977 EEE tests received some form of pass – which means that approximately 90% of the test applicants received some substantial benefit from the test.

Meanwhile, the follow-up study of the 1973 test group continues to accumulate data. In time this study should give definitive answers to the questions a large-scale equivalency testing program raises for a state university system – in particular to the question of how students gaining credit differ in their course plans and grades from those not gaining credit. The staffing, financial, and educational implications of this study are obviously important. Third-year findings are included below in Chapter V.

The 1978 English Equivalency Examination will be offered on April 29, 1978, and the essay will once again be scored by CSUC faculty in the middle of June. As this volume goes to press, discussions are underway about bringing the essay reading for the Spring 1978 English Placement Test to Pomona at or near the time for the EEE reading. Certainly, imaginative ways of bringing these complementary programs together are called for, most particularly in the light of the new CSUC graduation requirement demanding student writing proficiency.

1977 ENGLISH EQUIVALENCY EXAMINATION FACT SHEET

Test Date:	April 30, 1977
Test Used:	CLEP Subject Examination, Analysis and Interpretation of Literature (90-minutes) in combination with Two 45-minute essay questions prepared by CSUC English faculty (90-minutes).
Number of Candidates:	2,218
Number Passed:	884
Percentage Passed:	39.9%
Number of Semester Credit Units Earned:	5,304
Scoring Data:	
Minimum Passing Scores:	
Objective Part:	45 (CLEP scale) achieved by 58.6%.
Essay Part:	14 (out of a possible 24 points; equivalent to 43 on the CLEP scale) achieved by 59.5%.
Combined Score:	92 (49 objective + 43 essay, each on the CLEP scale) 41.5% achieved 49 or better on the objective test, and 59.5% achieved 43 or better on the essay test.
Mean and Standard Deviation:	
Objective Part	
Mean:	47.050
Standard Deviation:	9.119
Essay Part	
Mean:	44.804
Standard Deviation:	8.753
Combined Score	
Mean:	91.854
Standard Deviation:	15.616

Statistical Data

Correlation between CSUC Essay and CLEP Objective Tests:	.53
Reliability:	
CLEP Test:	.88
Essay Test:	.70 ± .11
Composite Score:	.865 ± .035

Essay Reading

Number of Readers:	36
Number of Colleges Represented:	All 19 CSUC campuses
Grading Scale:	6 point
Number of Readings per Test:	4 (2 independent readings for each question; additional readings to reconcile discrepant grades)
Total Essay Readings:	Approximately 10,000
Weighting and Scaling:	The essay score was converted by the equipercentile method to the CLEP scale (20-80), and the scores on both parts of the examination were added. Each part received equal weight.

COMPOSITE
ENGLISH EQUIVALENCY EXAMINATION SCORING DATA
1973 – 1977

	1973	1974	1975	1976	1977
Minimum Passing Scores					
Objective Part	45	45	45	45	45
Essay Part: Raw score on 0-24 scale (Scaled score on 20-80 scale)	13 (48)	13 (53)	13 (46)	13 (48)	14 (43)
Combined Score	100	105	99	101	92
Mean and Standard Deviation					
Objective Part					
Mean	47.448	49.444	47.180	48.995	47.060
Standard Deviation	8.854	9.001	9.069	9.425	9.119
Essay Part					
Mean	47.916	48.620	45.291	49.447	44.804
Standard Deviation	8.402	9.171	8.693	9.107	8.753
Combined Score					
Mean	95.363	98.064	92.472	98.442	91.864
Standard Deviation	14.869	15.865	15.402	16.456	15.616
Statistical Data					
Correlation Between CSUC Essay and CLEP Objective Tests	.4777	.5220	.5037	.5768	.53
Reliability					
CLEP Test	.88	.88	.88	.88	.88
Essay Test	.72 ± .18	.68 ± .15	.89 ± .14	.87 ± .04	.70 ± .11
Composite Score	.88 ± .05	.86 ± .05	.86 ± .04	.87 ± .04	.865 ± .035

COMPOSITE ENGLISH EQUIVALENCY EXAMINATION FACT SHEET
1973 – 1977

	1973	1974	1975	1976	1977
Number of Readers	75	49 —including 4 guest readers	44 —including 2 guest readers	32	32
Number of Colleges Represented	All 19 CSUC campuses	All 19 CSUC campuses and UCLA, UCSB, Stanford, Southern University in Baton Rouge, La.	All 19 CSUC campuses and UCLA and the University of Redlands	All 19 CSUC campuses	All 19 CSUC campuses
Number of Readings per Test	4	4	4	4	4
Total Essay Readings	Approximately 20,000	Approximately 20,000	Approximately 15,100	Approximately 14,400	Approximately 10,000

Weighting and Scaling

The essay score is converted by the equipercentile method to the CLEP scale (20-80), and the scores on both parts of the examination are added. Each part receives equal weight.

ADMINISTRATION OF THE TEST PROGRAM

Once again, Dr. Edward M. White, California State College, San Bernardino, served as director of the test program. Replacing Dr. Richard Lid as Associate Director was Dr. Rex Burbank, San Jose State University. Coordination of test administration took place through the test office of California State Polytechnic University, Pomona, under the direction of Dr. Robert Bradley. A new test manual was prepared by Betsy Barlow of the Educational Testing Service in Princeton, New Jersey, and the administration of the test went smoothly.

No changes were made in the existing statistical or computer operation, which allowed the pass and no pass letters to be ready for mailing to all students who took the examination within 24 hours of the end of the essay grading process. A second computer operation was prepared, however, under the supervision of Dr. Bradley, to allow for a second-level pass, for those students not receiving credit but performing well enough to merit exemption from the English Placement Test. (See Appendix C for a discussion of the relationship between the EEE and the EPT.)

This year the same procedures that were used to develop the essay test questions in 1976 were again adopted. The advances in essay test development that were made in 1975 and 1976 were an important contribution to the field of essay testing and we discovered once again that the time, effort, and money that were put into the creation of each essay question led to questions that were excellent for their purpose and that could be scored with accuracy. A committee based in northern California was under the leadership of Professor Rex Burbank, San Jose State University. The other in the southern part of the state was under the direction of Professor Eileen Lothamer, California State University, Long Beach. Working with Professor Burbank were Professor William Robinson, San Francisco State University; Professor Jim Frey, California State University, Fresno; Professor Jack Connor, California State University, Hayward. Working with Professor Lothamer were Professor Elaine Osio, California State University, Los Angeles; Professor John White, California State University, Fullerton; and Professor Agnes Yamada, California State University, Dominguez Hills. These committees developed essay questions during the course of the fall term and assigned them to freshmen composition sections on their own campuses for pretesting.

Following pretesting, grading sessions were held by the committees, careful assessment of scoring reliability and grade distributions for each proposed question then took place. Both questions were then revised and once again pretested. By the time the essay questions were ready for printing, the committees were confident that the questions would accurately reflect the students' abilities and that these questions contained no ambiguities or other problems that would lead to difficulty in grading. Results of the grading session in June demonstrated the value of the essay test development procedure. Once again, the questions worked very well, most particularly in allowing excellent students to show how well they could perform without distractions or ambiguities in the questions.

One important conclusion from this procedure is that any large-scale testing program must be prepared to spend both time and money on development of essay test questions. No matter how excellent a question may appear to be to a group of English professors, it may only be considered to be a good essay question after it has been pretested on real students and the resulting essays have been read carefully by trained readers. These readers can then judge whether a full and accurate range of scores is being elicited by that question and whether reliable scoring of the essays is possible. Certainly the essay test development and the essay scoring procedure for the English Equivalency Examination have developed into a highly reliable form of testing with important implications not only for testing but for the development of writing assignments in general. It has become clear to many of the teachers involved in the program that some essay assignments given to students in classroom situations lack the clarity, specificity, and ease of scoring that the English Equivalency Examination essay questions possess.

Finally, the California State University English Council continues to maintain close professional supervision over the entire test program. English Council committees evaluate the aims and procedures of the program and continue to support the test program by votes affirming the recommendations of these committees. (See *Comparison and Contrast*, 1975, p. 12.) The Council has suggested that campus English Departments use the fact that they sponsor the test to "improve relationships and communication between our departments and the high schools." Campus English departments continue to invite the program directors to speak at articulation meetings, and most of those meetings open new avenues of communication between the departments and local high school and community college English teachers. Only through continued participation by the English Council will the test continue to have the confidence of the English faculties throughout the system and continue to deserve that confidence.

III

THE ESSAY TEST

The following report was prepared by Dr. Eileen Lothamer, question leader for Question 1.

Summary Report on Question 1

For the 1977 English Equivalency Examination, Question 1 was designed and pretested by a committee of four state university English department faculty members (Elaine Osio, Los Angeles; John White, Fullerton; Agnes Yamada, Dominguez Hills; Eileen Lothamer, Long Beach). These four veteran readers, aware of the experience-based nature of the first question, met at California State University, Long Beach, for the first of three meetings on October 15, 1976, instructed to bring three such topics, classroom tested and scored prior to that meeting. During the first meetings, the committee reviewed the questions and scoring guides for the previous English Equivalency Examinations, checked the responses to the individual committee member's trial questions, and selected two trial topics to be pretested at each of the campuses and scored (with tentative six-point guides) before the next meeting.

At the second meeting, November 19, 1976, the group scored samples from the pretesting runs, chose the question that elicited the fullest range of response and scores, and developed an alternate form of the one question chosen to determine whether or not specific details should be incorporated in the writing assignment directions. The committee also worked on a scoring guide to be used and evaluated by all before the subsequent meeting.

At its third meeting, February 25, 1977, the members were asked to bring scored samples for both forms of the chosen pretested question and suggested revisions of the scoring guide based on evaluation of the papers themselves. Considerable time was spent on refinement of the scoring guide, a crucial evaluation instrument if reader reliability is to be achieved. Also, twenty-eight samples were selected for training of table leaders and readers on the opening day of the reading. Final wording of the question and scoring guide was established.

During the June 11 table leaders' meeting, eight samples were first scored without the scoring guide so that the table leaders might check the inductive process by which the Design Committee had arrived at the guide suggested for use. Discussion of the guide and clarification of training goals and methods followed, with a total of sixteen samples used to set a comparative frame of reference.

The afternoon session that day began with a surprise for the readers. They were asked to simulate the writing experience of the students on both examination questions but with only twenty-five minutes' writing time rather than forty-five. They were asked to respond seriously to the demands of the questions. Following that trauma, and a coffee break, the training process for readers began. First, careful attention was paid to the question and the three-part nature of the assignment: to describe what it would feel like to be reincarnated as a particular animal, to explain the choice of animal, and to determine how that animal would be like or unlike the student's present self. During that training session, readers scored eight samples of 2 and 3 and 4 and 5 and 6 level papers, with room discussion of the rationale of the scoring guide and emphasis on the necessity of using the guide to achieve consensus. Table leaders then instructed their respective tables on a quick overview of five booklets each, without scoring, to foster holistic reading and close scrutiny of the scoring guide with "live" material. Because of the readers' own writing experience, only fifteen minutes remained that day for reading, but they were primed with the memory of their own actual writing agonies and the comparison of samples and scoring guide.

The following morning and at each subsequent beginning of a morning or afternoon session, the question leader and table leaders (Ed Cairns, San Luis Obispo; Arlene Okerlund, San Jose; Larry Kramer, San Bernardino) conferred on samples to reiterate the concern with holistic reading and use of the scoring guide. In addition, table leaders began to select "live" samples for occasional group scoring at a table. As with previous English Equivalency Examinations, responses to Question 1 were scored by two readers, and, where discrepancies of two or more points occurred, by a third reader. Moreover, the table leaders provided continual spot checks of the scores of their readers. Table leader check sheets were submitted to the question leader who also checked papers to monitor tendencies of readers, table leaders (or the question leader) to settle at a 3/4 score or to score somewhat erratically without sufficient reference to the scoring guide. Such problems were refreshingly few after the first whole reading day. The table leaders and I

were satisfied that readers did score with careful reference to the scoring guide and that the test booklets received serious and close readings. The rewardingly low discrepancy rate pleased us all.

And again, although we did not expect perfection from even the best student papers, we were pleasantly surprised by the quality of the 5 and 6 level responses, given the impromptu nature of the writing and the forty-five minute time limit. The topic provoked response because all high school students could write without specialized knowledge and because the writing assignment, through its structure, aided the student in organizing a response and stimulated adequate development. But, as a result of bicentennial fervor or commercial influence, far too many students wanted to return as eagles. We all thought for a while of selecting only student "eagle" samples so that everyone could attest to our objectivity despite the repetition. Perhaps it had been chastening for the readers to remember their own choices during their "simulated" writing. Whatever the reasons, the readers once again proved to be dedicated and conscientious and the reading itself an occasion for esprit de corps. With fewer I-score papers and a generally higher score range overall, the examination might be showing increased self-selection by students, increased attention to writing skills in the schools, increased concern with the question design and the scoring guide. Many of the papers were obvious and literal, many witty or imaginative, developing the expository points implicitly. We hope that the following samples show a bit of the range we encountered.

Eileen Lothamer
June 26, 1977

Question 1. The following question was distributed to all students:

PLEASE READ AND THINK ABOUT THE FOLLOWING:

Sometimes it's fun to fantasize about reincarnation. One person might like to come back as a rhinoceros; another, as a butterfly, sea anemone, or fox terrier. If you were able to come back in your next existence as an animal, what animal would you choose to be? Why?

WRITING ASSIGNMENT:

- Describe what it would feel like to be that animal.
- Explain your choice.
- How would it be like or unlike your present self?

The following directions for scoring were distributed to all readers engaged in the grading of Question 1.

The student should be rewarded for what he does well in response to the question. Here the student is set a threefold task: to describe what it would feel like to be reincarnated as a particular animal, to explain the choice of animal, and to determine how that animal would be like or unlike the student's present self. The student is not required to have an accurate understanding of animal behavior.

Note that the question calls for description and analysis. However, in some papers the analysis may be implied through description, but such implication should be appropriate to the approach used by the student.

An extremely well-written response may be scored a point higher than it would be scored on the basis of content alone. A poorly written response may be scored a point lower.

Possible Scores:

6 A superior response will address itself to all aspects of the question. Though it may have occasional faults, it will be well-organized, detailed, and generally well-written.

5-4 These scores will be useful for a well-handled paper which is weak in some aspects of the superior response; e.g., it may slight one of the parts of the question; it may not be as clearly organized as the superior paper; it may have some minor grammatical inconsistencies. Otherwise, the paper should be competently written. Remember that a four is an upper-half paper.

3 This score will be useful for the following kinds of papers:

- those which are only descriptive;
- those which are only explanations of choice;
- those in which the language is overly clichéd;
- those which are overly repetitious;
- those which are general and superficial.

2 This score is to be used for papers which exhibit serious weaknesses in structure, syntax, diction, and/or development.

1 This score is to be used for papers which show very little understanding of the question or suggest incompetence in structure, syntax, and diction.

* Non-response papers or those which argue with or avoid the question should be given to the table leader.

The following student responses to Question 1 were sample papers used during the reading to illustrate the grades on the 6-point scale:

Score of One

If I were to return to this life I would surely return in the form of the majestic eagle. Unlike humans, an eagle's omnipotence and majestic power of flight is seldom challenged. The eagle is not bored by trivial tasks which most humans are forced to perform daily; instead, he is free to fly to great heights and receive the attention that is drawn by his handsome and streamlined body. I would fly high above humans to look down and think how foolishly and meaninglessly many of their lives are wasted.

Score of Two

If I was to come back as an animal I would choose to be an eagle. An eagle is a symbol of freedom, to do as it chooses and to glide above all other creatures as master of all. I enjoy my freedom, the ability to soar for hours and then dive, at speeds so fast that I am just a blurr in the sky, would exhilirate me. I be the master of the sky and to look down on the rest of the world would never seise to thrill me.

I am not as free as an eagle, yet I enjoy the freedom I have. To me freedom of the mind and body is what life is about and an eagle is just that, free.

Often I have imagined soaring in the sky, it is as if your body had no bounds, you are one with the world, you are one with the universe, yet you are separate from both and can take in all the beauty and splendor of it all.

As I see it an eagle is the perfection of freedom.

Score of Three

As I fantasize about reincarnation, I feel I would like to come back as a dolphin.

A dolphin is an intelligent and talented animal. If I were a dolpin in the Pacific Ocean, I would swim along the California coastline until I caught the eye of a talent scout for Marine World. Being taken captive, I would reside at Marine world; I would learn all kinds of tricks easily, because of my great intelligence, and amaze the people of the audiences.

I would choose to be a dolphin, a performing one that is, because I feel I would be doing something for a small part of the world. Although I don't spend much time swimming as a human, I'm sure I would as a dolphin. From my point of view being a dolphin would have some advantages. Can you imagine not having to cope with the anxieties and problems of the human race? All I would have to worry about is having water to swim and live in, not missing feeding time, and knowing my performance.

In a way, I can relate to this dolphin I have described; he would be bringing the pleasure of entertainment to people as a small part of me takes pride in doing. Just think, I could become another famouse dolphin like Flipper!

Score of Four

If I were able to come back in my next existence as an animal, I would choose to be a sea otter. I have always loved the ocean, and to be able to have the chance to be a part of it is something I fantasize about.

Being an otter would be being myself. Otters are fun-loving creatures with a tendency for being mischeivious. They are so full of life and vitality; always curious about new and different things. Fun-loving: yes . . . and certainly a bit mischeivious – that's me!

I can relate to an otter. I, too, am always exploring the unknown, and more than often finding myself in a jam because of it. But being yourself and trying to learn as much about the world around you as you can sometimes takes a little bit of getting into those "sticky situations."

Otters may be rascals at times; but they are also devoted parents. Have you ever seen an otter with her baby? The love and understanding she gives it is a beautiful sight to see. How do I relate to this? I build my life around children. I plan to be a Special Education teacher, and the patience and love it takes to work with handicapped and mentally retarded children is not easy. I have watched a mother otter teaching her baby the ways of the ocean. Baby otter may not always learn fast, but he tries hard, and mama is always there to encourage him.

It's like that with children too – especially those who have special problems. Encouragement, patience, and lots of love is what these children need the most. And it's not only mother otters that give their babies these basic needs, but other animals as well.

I think the final reason I would like to be an otter is because I love the water. During the summer I teach swimming to children. Being in the water is something I'm used to!

Yes, if I could be reincarnated – an otter would be what I'd choose. Full of life, mischeivious but also loving; that's an otter – and me, too!

Score of Five

If reincarnation does exist in this world, I would like to be reincarnated as an eagle. The glory that has been apportioned to this majestic bird by man as well as its position in the animal kingdom has affected my choice somewhat. I have always felt a solemn reverance for the eagle and have looked upon its

lifestyle as amiable. Its undisputable strength and regalness might parallel it's existence in the animal world with that of a great political leader or monarch of the human world.

It would seem that the eagle's impressive position would make it an animal of great assuredness. Yet, at the same time the eagle is not an extremely aggressive animal. It does not rule its kingdom with the same abominable force as that of the lion. The eagle rules its vast environment benevolently.

If I were to assume the role of the eagle and thus, be an eagle myself, I would soar through the sky and keep an ever present watch upon my kingdom. I would keep to my element as much as possible. The seasons would come and go and I would be a part of them all. I would feel the soft, rain massaging my body, the snow blanketing my wings, the wind beckoning me, and the sun, an unreachable god. Each day would be an event in itself, for there is always changes in weather. But, I would not limit myself to the day. At night, I would become a part of the heavens, I would be a winged astronomer and endeavor to reach each new star.

I would not be tied down by anything. Nothing would hold me to the earth and I would experience total freedom. I would avoid the world of man along with its hypocrisy and evil. To fly over his houses and myriads of highways would be enough for me. When I rested, I would choose not the tree planted by man but would stop only in the virgin wilderness.

I would perceive all that there is to perceive and would attain incomparable knowledge. I would learn what each acre of land was and know the movements of the sea. I would learn the secrets of time and space and yet not be governed by them.

All of this would be my dream of the perfect existence. It parallels very closely my ideas about freedom and wisdom and such at this time. But, I do not find many traces of the eagle in myself. I am governed by responsibility and bound by time and space. I am victim to materialism and futility.

Strength and wisdom are my goals, although I speak more of an inner strength than physical strength. I dream of doing something great one day, but one cannot build their life on dreams.

I feel an affinity with nature as does the eagle, but I find it hard to truly appreciate nature in a totally automated world. Thus, in consideration of the self-criticisms I have made, I see the eagle as my ideal self. I long for the freedom that winged animals seem to have found. There can never be such perfect freedom on the earth as long as man inhabits it.

Score of Six

Given a choice of what animal to reappear as in a future life, the choice is obvious – a flamingo.

Consider for a moment the existence of the flamingo, a communal creature standing stoicly in 6 inches of water all day somewhere in Africa. Rarely, if ever, is the flamingo subjected to the perils of noise pollution, smog. and urban crime.

If I could arrange the transformation or re-birth as a flamingo, while my life wouldn't be hectic, there would be daily events to keep my mind occupied. Occasionally (as we avid followers of Disney Nature movies note), a crusty alligator or crocodile will venture into the marsh and attempt to snap off a flamingo leg or two. Concurrent with the phrase "Flamingo dancing," all of that grating, stamping about is symbolic of a bird trying feverishly to save his legs from consumption.

The main activity of flamingo's however, is hunting down their food. They eat non-descript little fish, and peck them out of the water lightning quick, and gulp them down horribly. An avid admirer of sport myself (1974 pinball champion), this game of survival intrigues me, and makes me long for the life of the flamingo all the more.

Living as a flamingo, I would seize the opportunity to actively organize their ranks. For too many years the flamingo has been the silent majority in the bird world. All the other birds have great representation in the media; look at the field: "Jim Crow" symbolizes racism, storks symbolize sexual misconceptions, and toucans have cornered the market on fruit loops.

The treatment of the flamingo in this nation has been less than shabby. I mainly wish to become a flamingo, so as to develop a radical Flamingo Rights Organization. Calling it N.O.F. (National Organization of Flamingos), I would drive the apathy out of my fellow comrades. In my identity as a flamingo, I plan to capitalize on the color of flamingo Pink! After I mobilize Africa's (and America's) flamingos, when they are called "Pinko's," it will be a political commentary as well as a physical observation. N.O.F. fully intends to go "left-wing," flapping a revolutionary message all over the world – "Flamingo Power."

In addition to the opening field in Flamingo revolution, there are other attractive features about the flamingo. Most importantly, they are thin, willowy creatures, and as one who has been mildly obese from birth, the notion of continual svelte-ness appeals to me tremendously.

Probing further into my phyche; I find something in the flamingo's temperment that appeals to me. Wading in tranquil swamps, with only a few lapses in contentment for radical N.O.F. slogans and alligator attacks, is the kind of life that I find attractive. Since in this life, I have lacked both the religious insight, and sexual fortitude to become a nun. My only hope for true meditative contemplation is in my next life as a pink bird, somewhere in Africa.

**Recently, in a landmark decision, Judge Hiram Templin of Indiana granted damage payments to the N.O.F. for Steel and Plaster lawn ornaments figured after flamingos. In his decision, Templin called the statues, "grossly exploitive."*

The worst conceivable thing (for a flamingo) could happen however. I could be captured in a gunny sack and imprisoned in a U.S. zoo, forced to eat recently thawed shrimp with Red dye #2, and remain cooly collected when prepubescent children make remarks like: "Mommy what's wrong with that funny skinny bird? What happened to his other leg?"

The following report was prepared by Dr. Rex Burbank, question leader for Question 2.

Summary Report on Question 2

The development Committee for Question 2 consisted, as it did in 1976, of Professors Jack Conner, Hayward; James Frey, Fresno; and William Robinson, San Francisco. I served as chairman. All of us on the Committee are members of the English Departments at our respective CSUC campuses. We met first in October, 1976, at San Jose State University, where we selected six essay topics for pretesting from the sixteen (four by each person) we brought to the meeting. Professors Conner, Frey and Robinson agreed to have two topics tested at their campuses and bring the papers to the next meeting, late in November.

When we met again in November, it soon became clear that three of the topics pretested were unsatisfactory, first, because they did not yield responses that would allow student-examinees to show how well they could write; second, because they did not call upon the analytical skills Question 2 is expected to test; and finally, because they did not generate responses that would lend themselves to holistic scoring. Of the other three topics, none produced essays that would, in our judgment, permit readers to make discriminations on a six-point scale. We felt, however, that one of the topics could do the job if the directions were revised a bit. We changed the directions, making them more explicit, and pretested the topic again at Hayward. The revised version showed improvement, but the responses tended to be too general and philosophical and less analytical than they needed to be, so we revised the directions once more to correct that weakness and pretested the second revision at San Jose State. The third version proved to be most satisfactory. We devised the scoring guide and located samples to be used at the reading in June.

As was the case in previous years, Question 2 was designed to suggest an organizing principle upon which the examinee could write a 45-minute essay. It included two quotations, one from Kenneth Roberts' *Northwest Passage* and the other from *The Web and the Rock* by Thomas Wolfe. The assignment was to write an essay explaining the attitudes towards striving and succeeding in each of the passages, showing how the views differed, telling what weak or strong point each had. The suggested structure of the response was thus based upon definition or analysis, comparison, and contrast. With the organizing principle thus built into the assignment, the examinee could develop his/her thoughts within a ready-made framework. The rubric was constructed to enable readers to score all essays by the same qualitative standards.

The topic turned out to be a good one in two important respects, namely, that it enabled the examinees to write lengthy responses that balanced generalizations with details, and it tested the ability to deal with subtle, important distinctions in meaning.

Scoring of the essays required even more intense and prolonged concentration by the readers than in the past; we had to use more samples, and the table leaders (Professors Lennis Dunlap, Chico; David Rankin, Dominguez Hills; and Robert Wylder, Long Beach) and the chief reader applied especially intensive efforts in checking for accuracy. The results were gratifying. The question, we agreed, was a successful one, for it challenged the students, enabled the readers to discriminate according to the six-point scale, and provided a reliable indication of how well each writer could write on a difficult but fair topic.

On the first morning of the reading at Kellogg West, June 10, I met with the three table leaders and went through the samples selected by the development committee and related them to the scoring guide, or rubric. I asked the table leaders to read and score the samples again so we could check their scoring against that of the development committee. We found that second scorings were remarkably close to the first. We were thus ready for the afternoon training session with the readers, who gathered at 1:30. We had a total of 14 samples ready for use by the readers.

In the afternoon, the readers met with the table leaders and me and discussed the question, how it was developed, and the basis upon which it was to be scored. They read and scored samples of 2 and 3 and 5 and 6 papers, and I explained the rationale of the scoring and referred them to the rubric. They read and scored more samples and discussed them among themselves at the tables. At four o'clock I asked the table leaders if they felt their readers were ready to begin "live" readings; all three said yes, and so the reading began. At 4:45 p.m. three more samples were distributed and read, and the reading day ended at five o'clock with a review of the scoring standards.

All together about 30 samples were used during the reading. We used fewer the second than the first day, and fewer the third than the second. Table leaders spent all their time checking the scores of readers at their table. Readers were asked to adjust their scoring in accordance with the scores given the samples by the table leaders. In addition to sampling, table leaders read and scored the papers themselves, and recorded both their own and the readers' scores on a "check sheet." The table leaders, who read and scored the papers without looking at the scores assigned by the readers, brought the checklist with the test booklets to me, and I then read and scored them myself without looking at the scores given by the readers and table leaders. I was thus able to check continually the scoring of the readers and table leaders against my own. Where there were discrepancies, I could locate who was grading inaccurately and see that a correction was made; and where trends developed – such as a tendency to grade too high or too low or settle in the 3-4 range, the problem could be addressed immediately by further sampling in order to remind the readers to use the whole 6-point scale. Using these methods, we achieved a high degree of agreement between first and second readers, and in nearly 95% of cases the two readings produced scores that were within one point of each other. When a spread of two or more points occurred, the paper was read a third time, by one of the table leaders or by me. I am satisfied that overall the papers were graded with a very high degree of uniformity, reliability and validity.

In scoring the essays we did not expect perfection from even the best student responses, but for a score of six we required that the paper show a high degree of writing skill in doing what was asked by the assignment. In order to be given a six, a paper had to address itself to the issues raised by the two passages and explain what they meant, how they differed, and in what respects they were alike. The best papers were expected to demonstrate the ability to build on the implied organization with meaningful details, to develop the general points, and display a high degree of understanding of the quotations. The best essays showed an awareness that the Roberts' passage emphasized the

relationship between striving and succeeding, while the selection of Wolfe equated success with using one's talents. Readers understood and took into account the fact that students were required to read two rather sophisticated literary passages and then to write meaningful commentaries on them in a matter of only 45 minutes – a demanding task; nevertheless, in order to be awarded the highest score, examinees were expected to demonstrate evidence of ability to perceive the complexities implicit in the quotations and to relate them in clear, standard written English.

Rex Burbank
July 14, 1977

Question 2. The following question was distributed to all students:

> You have 45 minutes to write on the following topic:
>
> On every side of us are men who hunt perpetually for their personal Northwest Passage, too often sacrificing health, strength, and life itself to the search; and who shall say they are not happier in their vain but hopeful quest than wiser, duller folks who sit at home, venturing nothing and, with sour laughs, deriding the seekers for that fabled thoroughfare?
>
> Kenneth Roberts, *Northwest Passage* (1937)
>
> If a man has a talent and cannot use it, he has failed. If he has a talent and uses only half of it, he has partly failed. If he has a talent and learns somehow to use the whole of it, he has gloriously succeeded, and won a satisfaction and a triumph few men ever know.
>
> Thomas Wolfe, *The Web and the Rock* (1939)
>
> ASSIGNMENT:
>
> Write an essay on the two passages above in which you answer the following:
>
> – What attitudes do Roberts and Wolfe express towards striving and succeeding?
>
> – To what extent do the views in the Wolfe passage differ from those expressed in the passage by Roberts?
>
> – What strong or weak points does each position have?

The following directions for scoring were distributed to all readers engaged in the grading of Question 2:

The student is *asked* to write an essay which explains (1) the attitudes expressed in each of the passages, (2) the extent to which the views in the Wolfe passage differ from those in the passage by Roberts, and (3) the strong or weak points of each position. Papers should be scored for their *overall* quality, and the student should be rewarded for what she/he does well, but an essay that does not fulfill the three assignments either explicitly or implicitly should not ordinarily be given more than a 3. The topic calls for comparison and contrast. Upper-half papers will recognize that the passages are comparable in what they say about the desirability of acting upon aspirations or using one's talents, but they will also point to the distinction to be made between Wolfe's emphasis upon the value of succeeding and Roberts' upon the value of the quest itself.

Papers that reveal an understanding of the metaphor in the Roberts passage or a consciousness of tone may be given an extra point. An especially well-written essay may be scored a point higher than it would be scored on the basis of content alone; a poorly written response may be scored a point lower.

Possible Scores:

6 Although a 6 paper does not have to be perfect, a superior response will be a well-organized essay that (a) addresses all three aspects of the assignment; (b) reveals a clear understanding of both passages; and (c) demonstrates a high degree of competence in writing.

5 Essays given a score of 5 may emphasize one or two of the three assigned questions. Although the 5 response will be less fully developed, less well organized, and perhaps less fluent than the 6 response, and may have minor flaws in grammar and diction, it will show clear understanding of the two passages and overall competence in writing.

4 Although a 4 paper may have flaws in grammar and diction, show less fluency than a 5 or 6 paper, misinterpret one of the two passages, or deal with only a part rather than all of the assignment, it will display general overall competence.

3 Although a 3 paper may be argumentative or personal, may show flaws in sentence structure and diction, be superficial in content, or include some irrelevancies, it will show understanding of at least one of the two passages and display evidence of organization.

2 A 2 paper will show serious misunderstanding of one or both passages *and* be poorly written.

1 This score should be assigned to papers that display little understanding of the question or the passages, and show *incompetence* in structure, usage, and idiom.

0 Non-response papers and those completely off the topic should be given to the table leader.

The following student responses to Question 2 were sample papers used during the reading to illustrate the grades on the 6-point scale:

Score of One

In the passages both men put emphasis towards the pleasures and fulfillment in having personal goals in life. Also anyone having a chance to fulfill these goals and not taking advantage of it have in life have failed. But someone who makes the sacrafice and reaches his goals accomplishes something few men do.

In the Wolfe passage the idea of fulfillment of personal goals is dependent oneself and failure to reach these goals does not make a lesser man; Unlike the Roberts passage where he criticizes those who don't have any courage to challenge the unknown. Roberts

Score of Two

Many talents are some times wasted, because people don't wish to put the time and effort into worthwile projects. Generally speaking, people are lazy. Its hard to find something to be tottaly interested in, and dedicated to. For this reason, people have many different attitudes about striving and succeeding.

In the Northwest Passage by K. Roberts the attitude toward striving and succeeding is that health, strength, and life itself are too often lost. To have a goal out of life is a good idea, but don't kill yourself accomplishing it. Having a goal does not necessarily make you happy either. To be content with what you do and what you have can be a happier life than constantly striving. In another passage, The Web and the Rock, by Thomas Wolfe, the attitude is to make use of your talents to the best of your ability. Not using talents one is born with is a failure in life, but knowing your talent and putting it to use is succeeding. Both passages seem to have different views.

In the first passage by K. Roberts it seems to say do what will make you happiest. If you must conquer something or complete a goal, or always be striving for more out of life than you have, then do it. But if your happy with what you have, or just sitting at home watching T.V., then do it. Just sit back, relax and enjoy, which was the view from the second passage by Wolfe. Out of each passage both strong and weak points were found.

The strong point of the first passage was don't sacrafice health, strength, and life in your quest. Maybe go after something that is in reach, not something impossible. The strong point from the second passage was not to be a failure in life. The weak point being to learn your talents and use them.

So find your talent, and learn its use, so you won't be another failure in life sitting at home watching T.V.

Score of Three

In the two passages by Roberts and Wolfe, both of the authors write about sucess and failure. Both of the authors treat the subject in a different manner. Roberts gives examples to illustrate his point and Wolfe just states his viewpoint.

Roberts is saying that the adventurer is not necessarily happier than the folks who sit at home. They both are happy in their own way. The adventurer who strives for their own personal success and hopefully finds it. And then there is the man who sits at home, who Roberts calls "wise", he can not be less happy than the adventurer for his is striving for his own type of success. Roberts writes that each man has his own quest and if he has found it, than he succeeded.

This is contrary to the viewpoint of Wolfe, however. Wolfe's view is that if a man has a talent and does not use it to its full capacity, he has failed. He says that only the man who learns how to use the whole of it is triumphant in their quest for success. Wolfe does not think that a man could be a success without using his talent to its' potential.

Roberts feels that each man sets his own goals and when he reaches them, he has succeeded. Wolfe, on the other hand, feels that one must use his talent to the fullest to become successful. Wolfe argues that few men succeed, whereas Roberts thinks that happiness is success.

Both authors stand on different viewpoints, although, talking about the same subject. In Robert's passage, he states examples which help to clarify his attitude towards success. He says that we either look for our personal quest or we find our happiness in ourselves. He speaks of happiness and success in the same manner, i.e. man's happiness in his success. Wolfe, though, states his viewpoint in another manner, comming directly to the point. He states it as a fact without backing it up with an example Wolfe speaks strictly of talent, whereas, Roberts speaks of happiness.

I tend to agree more with Roberts because I also believe that if a man is happy, he has found his success. In Wolfe's passage, he says a man can only be happy if he uses his talent to the fullest. If he had no talent, does it mean he has failed. If he did have a talent and used it to its fullest, can he say he succeeded?

Score of Four

Roberts and Wolfe: Striving and Success

Kenneth Roberts and Thomas Wolfe convey similar attitudes toward striving and suceeding in the two passages. The authors however, although they express a similar message, also have some differences. In the presentation of their messages, both men have weak and strong points.

Both Roberts and Wolfe believe that striving is an important part of a man's life. The authors agree that pursuing ones own dreams and goals is a great attribute to a persons happiness.

Although the two men agree that striving for one's ideals is beneficial to man, they also have differences in opinion in their theories. Roberts states that the hunt for one's own 'Northwest Passage' is in vain. This author does not believe that man can reach his goals in his lifetime. Wolfe, on the other hand, believes that the happiness man feels, comes from, not the quest itself, but the satisfaction of success.

Thomas Wolfes theory is that, men who develop their talent to its full extent will succeed with a greater triumph than men who use only a fraction of their abilities. The positions which Kenneth and Wolfe take have both strong and weak points. The strong point in the two writers' positions is that they are encouraging people to live to its fullest extent, reach for the highest goal, and develop & use their own talents to the greatest capacity. The two authors leave men with a sense of encouragement that everyone can succeed as long as he tries to the best of his ability. Both authors have weak points, however, particularly Thomas Wolfe. In his passage, Wolfe makes no reference to those men who believe they have no talents. The author begins every line with 'If a man has a talent . . . . ' Men, therefore read the passage and tend to believe that not every man has a talent. These men may believe that they themselves are not endowed with any talent and therefore, not able to experience the great sense of achievement which Wolfe mentions.

Although the two passages are written by two different men, they convey the same meaning. Both passages convey the importance of striving for ones goal or, as Kenneth Roberts states, ones personal Northwest Passage. Although the two writers convey the same message, they have some differences. The main difference is that Kenneth believes that happiness results from the search, while Wolfe states that satisfaction is a result of success. Although the passages are short and meaningful, they have both weak and strong points. Both authors strong points are similar to each other, while Thomas Wolfes weak points alone stand out.

Score of Five

What is Success?

We like to succeed in what we do. But our success depends not only upon our own perserverance but also on what it is we are doing. Success means different things to different people. Is it the effort made to reach a goal or the actual attainment of that goal that constitutes success? Kenneth Roberts in his book Northwest Passage and Thomas Wolfe in The Web and the Rock express differing opinions.

Roberts states that the quest in itself is enough, that dedication and perserverance make a man's effort worthwhile. Wolfe believes that effort alone is not success. In fact, it is a total waste, a mis-use of talent. He believes that success is the efficient use of one's talent and anything less is failure.

Wolfe's views differ greatly from Robert's at first glance. They are not actually that different. They both believe that success is doing what one finds satisfaction in. The two men simply differ on the definition of "satisfaction."

Both views have good and bad points. Roberts expresses a "frontier spirit" in this passage; a more spiritual rather than material quest. But this spirit does not build nations or economies. Wolfe speaks of a man's talent, his spirit as a resource to be efficiently allocated. It is an impersonal view but can be very profitable in a material sense.

Success is satisfaction in trying or achieving depending on our own personal goals. What a dull world it would be if we were all "frontier" men or all "economic" men.

Score of Six

"A man's reach should exceed his grasp, or what's a heaven for?" Everyday, in every field of endeavor, human beings seek to attain goals which are beyond their immediate capabilities. Some of these goals are seen by others as hopeless fantasies; most, however, are simple goals which, given time, motivation, and diligence, will be reached. It is the "hopeless fantasies" which Kenneth Roberts addresses in his excerpt, while Thomas Wolfe speaks to the matter of reaching one's realistic goals.

Being an athlete, to wit, a swimmer, I can most easily apply the two passages to the world of athletics. I find that Roberts' view of personal ambition is the more easily digestible of the two, because while Wolfe says that anything less than 100% in every area is at least partial failure, Roberts asks who is to say

such a thing. Wolfe makes no allowance for personal preference if one has a certain talent; one must, he says, utilize one's talents or one is a failure. Roberts, on the other hand, seems to admire those whose reach exceeds their grasp, but does not castigate those "wiser" folks who decide to play it safe.

In fact, I think I would admire more the person who struggles valiantly (though perhaps uselessly) to attain a goal than one whose natural talent, while being used 100%, is not being provided with a challenge. It happens all the time in swimming: a youngster who, at age 10 or 12, has so much natural ability that he's breaking every record in sight, quits when in high school because the sport has lost its challenge. Meanwhile, the "second-stringer" who's struggled along for seven years suddenly blossoms at the high school level. There must be a challenge to one's ability, so that there is always something to strive for. Once you've reached the top, there's only one way to go, and it isn't up.

Roberts says that "too often", men sacrafice their health in a struggle to achieve. I see nothing wrong with such sacrifice if it is what one truly wants and if no one else is harmed by it. As Wolfe says, if one is using his abilities 100%, even if not "succeeding," he gains a personal triumph and satisfaction few will ever know, a feeling which more than compensated for any loss in physical health.

Where would the world be if men (and women) throughout history had not striven for goals seemingly unattainable? Neil Armstrong's first steps on the moon were probably the greatest single example of a once "impossible dream" being realized. World records in sports are constantly being broken, simply because humans refuse to believe that they can do no better. It is only by seeking to improve ourselves, by trying to find our own Northwest Passage, that sanity can be maintained and a purpose to life be found.

Success is in the trying, not in the succeeding.

TABLE CHART/READER ROSTER

June Essay Reading 1977

Question 1

Dr. Eileen Lothamer
Question Leader

Table 1

Edward Cairns
CPSU, San Luis Obispo
Agnes Yamada
CSU, Dominguez Hills
John B. Dalsant
Humboldt State University
Stanley Cook
CSPU, Pomona
Sam Bullen
Sonoma State College
Frank Hubbard
CSU, Sacramento

Table 2

Arlene Okerlund
San Jose State University
Joanne Lynn
CSU, Fullerton
Robert Hodgman
CSU, Los Angeles
Donald Keesey
San Jose State University
William Phillips
CSC, Stanislaus
David Andersen
CSU, Northridge

Table 3

Larry Kramer
CSC, San Bernardino
Jill Cohn
CSC, Bakersfield
Patricia Brenner
CPSU, San Luis Obispo
John White
CSU, Fullerton
Clark Brown
CSU, Chico
Shelby Steele
San Jose State University

Question 2

Dr. Rex Burbank
Question Leader

Table 4

Lennis Dunlap
CSU, Chico
William Cowling
CSU, Fresno
Solomon Iyasere
CSC, Bakersfield
Dan Buerger
San Jose State University
Michael Holland
CSU, Fullerton
Ruth Brown
San Diego State University

Table 5

David Rankin
CSU, Dominguez Hills
Margaret Doane
CSC, San Bernardino
Jack Conner
CSU, Hayward
Edward Borowiec
CSU, Long Beach
Alice Brekke
CSU, Long Beach
Arthur Lane
CSU, Northridge

Table 6

Robert Wylder
CSU, Long Beach
James Kohn
San Francisco State University
Tony Arthur
CSU, Northridge
Betty Hanson-Smith
CSU, Sacramento
Robert McDonnell
CPSU, San Luis Obispo
Katherine Carter
CSU, Los Angeles

IV

STATISTICAL REPORT

by Robert Bradley[1]

October 1977

The Test.

The fifth administration of the California State University English Equivalency Test was conducted on Saturday, April 30, 1977.[2] The test has two ninety-minute sections; the first is an objective, multiple-choice part, and the second is an essay part with two questions. The "Analysis and Interpretation of Literature" Subject Examination of the College-Level Examination Program (CLEP) was used as the objective part of the test. The two questions of the essay section were developed and scored by the English professors of The California State University and Colleges.

Scores.

1. **Objective Score.** The objective part of the test was scored by the Educational Testing Service, and scores were reported on the CLEP Subject Examination scale, which ranges from 20 to 80.

2. **Essay Reading Scores.** Each of the two essay questions was graded by two readers, and scores were given on a scale of 1 to 6. For a single reading of a question, a score of 4 or higher represented a passing performance. A zero indicated that the student did not attempt to answer the question.

3. **Essay Question Score.** The scores for the first and second readings of each question were combined to produce a question score, which ranged from 0 to 12.

4. **Raw Essay Total Score.** The two question scores were added to produce a raw essay total score. This score fell between the limits of 0 and 24.

5. **Converted Essay Score.** The raw essay total scores were transformed to the same scale as the objective scores to permit the addition of objective and essay scores in a meaningful way.

6. **Composite Score.** The objective score and the converted essay score for each person were added to produce a composite score. Since both the objective and the converted essay scores are expressed on a 20 to 80 scale, the composite scores ranged from 40 to 160.

Candidate Groups and Norms Samples.

1. **Total Group.** The 2,218 students who took both sections of the English Equivalency Test constituted the total group. The objective part of the test was taken by 2,223 students, but 5 of these students failed to take the essay section.

---

[1]I wish to thank Dr. John Bianchini, Head, Statistical Research, ETS Berkeley, for his substantial contributions to prior editions of this report and for his critical reading of a draft of this edition.

[2]See W. Cowell, "The California State University English Equivalency Test, May, 1973," Statistical Report SR-74-10 (Princeton, N.J.: Educational Testing Service, February 1974). See also R. Bradley, and J. Bianchini, "Statistical Report," in E.M. White, *Comparison and Contrast* 1976 (Long Beach: Office of the Chancellor, The California State University and Colleges, 1976).

2. **Equating Sample.** The equating sample consisted of the first 531 papers scored at the essay reading session. All of the essay books from the various test centers were systematically scrambled prior to the reading to insure that they would be read in an essentially random order. Inasmuch as the means and standard deviations of both groups were not significantly different, the equating sample was assumed to be representative of the total group.

OBJECTIVE SCORES

	Mean	S. D.
Total Group	47.05	9.12
Equating Sample	47.51	9.21

The data from the equating sample were used to equate (transform) the essay scores to the objective score scale. A sample was used instead of the total group so that computer processing and statistical analysis could be done as the reading session progressed.

3. **CLEP Norms Sample.** This group consisted of 541 college students who participated in the 1964 norming administration of the CLEP Analysis and Interpretation of Literature examination.

4. **Texas Norms Sample.** This sample consisted of the 188 students included in the 1970 validity study of the CLEP examination at the University of Texas. Scores on the CLEP test were correlated with final grades in relevant courses.

5. **CSUC Norms Samples.** The essay questions of the English Equivalency Test were administered to six different groups of California State University students. These essay papers, indistinguishable from regular papers, were interspersed with the candidate essays and scored during the reading session. The average essay scores for students who earned various letter grades in English composition courses and related courses were determined, and the information was used to help establish the minimum passing score for the essay part of the examination.

Three of the norms samples came from California State University, Fullerton, and three samples were drawn from California State University, Long Beach. The largest of the Fullerton samples contained 31 students who had just completed Communications 103, a course in applied writing for non-communications majors. The second Fullerton sample consisted of 12 students in Communications 101, Communications Writing, and 9 students in Communications 102, the second semester of 101. (The first and second semester samples were combined because of their small size.) Communications 101 satisfies the university's English composition requirement. The third Fullerton sample consisted of 9 students in English 103, Seminar in Writing, and 13 students in English 105, Introduction to Creative Writing. Both of these courses satisfy the English composition requirement.

One sample from Long Beach contained fifteen students who had completed English 184, Composition and Literature, a second semester composition course required of all English majors. Two of the Long Beach samples consisted of upper division students. Norm papers were written by seventeen students in English 300, Advanced Composition, which is required of all single subject credential candidates who either do not take or fail to pass the Advanced Writing Test. Papers were also submitted by 25 students in English 310, an applied composition course required of all credential candidates in English. One-half of the class time in English 310 is devoted to the principles of teaching composition. For each norms sample, the essay questions were administered to the students during the last several weeks of study designated for the course. The data from the CSUC norms samples are summarized in Tables 1-6.

6. **CSUC English Professors Sample.** Prior to the beginning of the essay scoring session, both essay questions were administered to the thirty-one CSUC English professors who scored the candidates' papers. Fictitious names were used, so the readers' papers could not be distinguished from student papers. The professors were given only twenty-five minutes to write each essay question, rather than the forty-five minutes allotted to the candidates. The mean total essay score for the thirty-one readers was 19.68, the standard deviation was 3.23, and the range of scores was 11-24.

The Equating Procedure.

Scores which are expressed on different scales cannot be combined without giving a greater weight to the scores on the scale which has the largest range of numerical values. The raw essay total score scale extends from 0 to 24, while the objective score scale extends from 20 to 80. If scores from these different scales were combined, the objective score would influence the composite score about three times as much as would the raw essay score. A solution to this problem is found by transforming the units of one of the scales so they are numerically equivalent to those of the other scale. This transformation can be performed so that the "equated scale units" represent equal increments in *relative position* on each of the score scales. Since the objective scores were reported on the CLEP scale, the essay scores were transformed to the CLEP scale by means of the equipercentile method of equating. In so doing, a given converted score represents the same percentile rank on both the essay score distribution and the objective score distribution. In other words, equal converted scores on the two parts of the English Equivalency Test represent the same *level* of performance on each part.

Technical Notes on Equating.

1. In the strictest sense, the score transformation procedure was employed to establish scores that are comparable, rather than truly equivalent. The terms "equating" and "equivalent scores" are fully appropriate only when the two tests are parallel forms designed to measure the same psychological function. Because the objective and essay sections of the English Equivalency Test do not measure the same psychological function (see Section X), the score transformation technique produces scores which are comparable but not equivalent. However, the mechanics of transformation are exactly the same for producing either comparable or equivalent scores, and the only difference is one in interpretation. For convenience, the term "equating" is used rather than a technically correct phrase, such as "establishing comparable scores."[3]

2. Two basic methods of equating are available, and when the score distributions are very similar in shape, these methods yield almost identical results. Given similar score distributions, the linear method, which defines equivalent scores as those represented by equal standard-score deviates, is preferable. The linear method is completely analytical (determined by a mathematical relationship) and is free from judgments of curve smoothing. However, if the two score distributions differ in shape, the linear method fails to adjust for this difference.[4] Thus, when the distributions have different shapes, or when the shape of one or both distributions is not known or cannot be accurately predicted, the curvilinear or equipercentile method is to be preferred. If the two score distributions differ in shape, the equipercentile method will stretch or compress the score scale of the distribution being converted to fit the shape of the other distribution. This method defines equivalent scores as those scores which represent the same percentile rank on both distributions. Using the score conversion table (Table 7) for illustration, a raw essay score of 14 corresponds to a converted score of 43. The percent of students scoring below 14 on the raw essay scale is the same as the percent scoring below 43 on the objective scale. Thus, in the sense that the two scores represent the same relative rank (percentile) within their distributions, they are equivalent scores.

[3] W. H. Angoff, "Scales, Norms, and Equivalent Scores," in R. L. Thorndike (ed.), *Educational Measurement*, 2nd. ed. (Washington, D.C.: American Council on Education, 1971), pp. 562-568; 590-597.

[4] *Ibid.*, p. 564.

3. Percentile rank values for each point on the objective score scale were plotted on normal probability graph paper. These points were connected by a smoothed curve, and for the points on the curve corresponding to the percentile rank of each essay score the objective score equivalent was determined.[5]

4. If the two score distributions differ in shape, the use of the equipercentile equating method, as noted above, will modify the essay raw score scale by altering the relative distances of some scores from the median. This modification of the raw score scale will result in a less-than-perfect correlation between raw scores and converted scores. Another factor can also contribute to an imperfect correlation. The equating process is based on a sample of essay scores, but the conversion process is applied to the total population of essay scores. If the shape of the total essay distribution differs slightly from the shape of the sample essay distribution, additional discrepancy may result between the raw score distribution and the converted score distribution. The correlation between raw essay and converted essay scores for the 1977 English Equivalency Test was .9904, which was slightly lower than in prior years. The correlations for the years 1974 through 1976 were .9993, .9955, and .9967. These relatively high correlations (.99+) indicate that the linear equating method could have been used without inducing serious discrepancies between the percentile ranks of equal scores on the converted essay and objective score scales. And, if the converted essay passing score is also derived by the linear method, the effect of discrepant percentile ranks between equal scores is virtually eliminated as far as passing rates are concerned. For example, if in 1977 the linear equating method had been used with a linearly derived converted essay criterion, only fourteen students (.7% of the total group) would have dropped from the pass category to the placement category.

Passing Scores.

1. **Basic Passing Scores.** An examinee was required to achieve a certain score on both the essay and objective parts of the examination in order to pass. These passing scores were added to arrive at a minimum composite passing score. Data from the CLEP norms sample were used to establish the passing score for the objective part. Since for both of these samples the average CLEP score for C students was 49, a score of 49 was set as the passing requirement on the objective part.

   For the essay part of the examination, a score of 4 represented acceptable performance for given essay reading. The criteria by which a reader was to distinguish between acceptable (4 or higher) and unacceptable (3 or lower) performance were carefully delineated in training sessions prior to the reading. These criteria represented a consensus among leading State University English professors. Each essay paper had two questions which were given two readings each, for a total of four reading scores. A score of 16 (on the 0 to 24 essay total score scale) thus represents an average score of 4 on each of the four readings. However, inspection of the frequency distribution of essay total scores reveals that only 29.3% scored 16 or higher, indicating that 16 would be a rather rigorous passing score. Also, setting the passing score at 16 would disallow any consideration for error in the scoring process. On the other hand, an essay score of 12 means that the average reading score was 3, or that on the average not one of the four readers found the performance on either of the two questions to be acceptable. A compromise was reached with a raw score of 14, which means that on the average two of the four readers judged the performance to be acceptable. The basic passing score on the essay part of the exam was therefore tentatively set at 14.

   The data from the CSUC norms samples (Tables 1-6) were also evaluated in an attempt to set an appropriate and defensible passing score for the essay part of the examination. The basic premise was that the essay passing score should be at least as high as the average essay score of students who earned grades of C in freshman English in the CSUC norms samples. For the Long Beach English 184 sample, the average essay score of C students was 13.33. However, inasmuch as none of the samples was drawn from a standard first semester English composition course, and because there were so few C students in any of the norms samples, it did not

[5] *Ibid.*, pp. 516, 571.

seem appropriate to use the scores of C students as a standard. It could be argued that a defensible passing score would be the average score achieved by all lower division norms sample students who earned a grade of C or higher (A + B + C). The Long Beach English 184 sample of this value was 15.07. For the combined Communications samples (101, 102, 103) at Fullerton the average score of all students earning C or higher was 13.81. The corresponding value for the Fullerton English 103 and 105 samples was 12.45. For the combined upper division samples (English 300 and 310) at Long Beach the value was 15.13. Thus the lower division "C and higher" averages ranged from 12.45 to 15.07, and the upper division average was 15.13.

In addition, it should be noted that grades in freshman English classes may be affected by factors essentially unrelated to writing ability, such as attendance, promptness with assignments, attitudes, and motivation. Given these considerations, the grades earned by students in the norms samples were used only as supplemental evidence in setting the passing score, with primary weight being given to the essay scoring standard developed by the CSUC English professors.

From the data presented in Tables 1-6 it is apparent that a passing score of 14 represents a level of competence slightly higher than that achieved by A + B + C students in all of the norms samples except the Long Beach English 310 group and the Fullerton Communications 101 and 102 group. The decision to set the basic passing score at 14 represented a consensus among the project directors, chief readers, and consulting statisticians that this score was solidly supported primarily by the carefully developed standards of the CSUC English professors, and secondarily by the performance data in the norms samples.

2. **Compensation Model.** Although in prior years a limited amount of compensation was permitted on both essay and objective scores, this year compensation was restricted to objective scores only. That is, an objective score could drop below the passing level if the essay score was sufficiently high. Since the basic objective passing score was set at 49 and the essay passing score was set at 14 (which converted to 43), the minimum composite passing score was set at 92 (43 + 49). The limit within which compensation would operate for objective scores was determined by defining an absolute minimum, below which no compensation on the essay could be justified. The objective minimum was set at 45, because this score minimized the number of A, B, and C students who would fail and the number of D and F students who would pass in both the CLEP norms sample and the Texas norms sample. To summarize, objective scores were allowed to drop as low as 45, provided they were accompanied by an essay score of at least 47 (15), thus resulting in a composite score of 92.

Placement Scores.

This year a middle score category called "pass for placement" was created. Students who did not pass for credit but who scored above a lower set of criteria were granted a recommendation for placement in a standard English composition course, and were also exempted from the English Placement Test (EPT). Students whose scores failed to meet these lower criteria received no recommendation and were required to take the EPT.

A simple conjunctive model was used for the "pass for placement" criteria. That is, no compensation was permitted, and a student had to achieve a certain minimum on both parts of the test. The essay score required for placement was set at 11 (34). A score of 11 can be represented by an average score of 2.75 on each of the four readings, or by three reading scores of 3 and one of 2. The project directors felt that a student who performed at this level exhibited the basic competence necessary for acceptable performance in a standard freshman composition class. The minimum objective score for placement was set at 35. This score represented a somewhat-better-than-chance performance on the multiple-choice part of the test, and was indicative of at least minimal competence in reading interpretation.

In addition, all persons who received objective scores below 35 but whose essay scores were 15 (47) or higher were included in the "pass for placement" group. Such a relatively high performance on the essay would argue for the student's readiness for a composition course, regardless of the low objective score. Fifteen students were added to the placement group on this basis. One student who scored 10 on the essay was also added to the placement group on the strength of an objective score of 61, which indicated considerable strength in reading comprehension.

Pass Rates.

1. **Objective, Essay, and Total Pass Rates.** A total of 2,218 students took both parts of the English Equivalency Test. The 884 students who passed the exam qualified to receive six semester units of credit in freshman English. The percentage of the total group which passed each part of the exam, using the basic passing criteria of 14 on the essay part, is given below. The percentage passing the total exam without compensation (requiring 14 on the essay *and* 49 on the objective part) and the percentage passing with compensation (as defined in Section VI, 2, above) are also listed.

	Objective Test		Essay		Total Exam			
					Without Compensation		With Compensation	
	No.	%	No.	%	No.	%	No.	%
PASS	921	41.5	1,320	59.5	714	32.2	884	39.9
NO PASS	1,297	58.5	898	40.5	1,504	67.8	1,334	60.1

It is evident from these data that a score of 14 on the essay represents a somewhat less difficult passing criterion than a score of 49 on the objective part. Accordingly, while only 22.4% of those students who passed the objective part failed the essay, 45.8% of those students who passed the essay failed the objective part. Expressed as percentages of the total group, 9.3% passed the objective test and failed the essay, while 27.3% passed the essay but failed the objective part. In 1976, 18.3% of the total group passed the essay but failed the objective test, and only 12.2% passed the objective part and failed the essay. These figures point to a sharp reversal in the relative difficulty of the pass criteria for the two parts of the test. In 1975 and 1976, the pass criterion on the essay was slightly more rigorous than on the objective part, even though some essay scores of 13, when accompanied by high objective scores, were allowed to pass. In 1977, although no essay scores of 13 were allowed to pass, the pass criterion on the essay became relatively much less difficult than on the objective portion, due to a large increase in the percentage (from 46.0 in 1976 to 59.5 in 1977) of students who scored 14 or higher on the essay.

2. **The Effect of Compensation.** The effect of the compensatory model was to raise the total pass rate from 32.2% to 39.9%. The compensation standard permitted one hundred seventy students with objective scores of 45 through 48 and essay scores of 15 or higher to pass the examination.

Technical Notes on Differences Among 1975, 1976, and 1977 Data.

1. **Differences in Mean Objective Scores.** The average objective score in 1977 (47.05) was 1.87 points lower than in 1976 (48.92), but only .13 points lower than in 1975 (47.18). The fact that the CLEP mean scores in 1975 and 1976 were substantially lower than in 1977 is most likely a reflection of the fact that alternate forms of the CLEP test are administered in successive years. Although the two forms have been equated, converted scores on one form of the test probably do not represent exactly the same level of performance as identical scores on the other form of the test. This inference is based on the likelihood that the two forms may differentially emphasize the measurement of several complex intellectual functions. Thus the conversion resulting from an equating study would not be totally independent of the individuals from whom the data were drawn in developing the conversion, and such a conversion would not necessarily be applicable to all student populations. In the present case, it may be entirely possible that the higher performance of the 1976 group is due simply to the fact that the group took the easier form of the CLEP test, since the relatively small difference in the CLEP mean scores for the 1975 and 1977 groups that took the same form of the test clearly implies a consistent level of competence from year to year.

2. **Changes in Converted Score Equivalents.** For the 1977 data, converted score equivalents for essay scores were substantially lower than those obtained in 1976. For example, in the critical pass-fail range, essay scores of 13, 14, and 15 converted to 48, 52, and 55 in 1976, while in 1977 the same essay scores converted to 40, 43, and 47. Factors contributing to these differences in converted score equivalents include the changes in mean

objective scores noted above, and year-to-year differences in the mean raw essay score. If the essay performance remains constant, a drop in objective scores results in lower converted score equivalents. Thus the effect of the lower objective scores in 1977 was to lower the converted score equivalents. The primary reason, however, for the drop in converted score equivalents was the relatively large increase in the mean essay score. To illustrate this factor, in 1977 the mean raw essay score (14.21) was .9 raw score points higher than in 1976 (13.31). This rise in the essay scores means that a given essay score in 1977 represented a lower percentile than in 1976, and thus converted to a lower equivalent score on the CLEP (objective) scale.

3. **Discrepancies Between Mean Converted Essay Scores and Mean Objective Scores.** The equipercentile equating process shapes the converted score distribution to exactly fit the objective score distribution, and thus the medians and means of both distributions should be equal. There are, however, several factors which have the effect of producing slight discrepancies between the mean converted essay score and the mean objective score. The amount of discrepancy between these two mean scores will vary each year as a result of variations in the degree each factor is present, and variations in the combined effect of the individual factors, which operate to either reinforce or cancel the effect of other factors. To illustrate this variation, in 1975 the discrepancy was 1.9 score points (mean objective = 47.2, mean converted essay = 45.3). In 1976, the mean converted essay score (49.5) was .5 points higher than the mean objective score (49.0). However, in 1977 the mean converted essay score (44.8) dropped over 2 points below the mean objective score (47.1).

   The factors producing discrepancy stem from two sources: 1) The fact that the equating is performed with a sample, rather than with the total group. If the equating sample score distributions (objective, essay, or both) differ from those of the total group with respect to either the shapes of the distributions or the values of the medians, discrepancy will result; and 2) the processes of rounding and curve smoothing in the equating procedure, which result in small differences between the mean converted essay score and the mean objective score.

Correlations Between Readings and Between Questions.

The correlation between the scores for the first and second readings of Question 1 was .51, and for Question 2 was .61. These figures, when expanded by the Spearman-Brown formula, represent the reliability of the grading procedure, and are discussed in the section on reliability. The scores for the first reading of Question 1 correlated .32 with the scores for the first reading of Question 2, while the scores for the second reading of both questions correlated .33.

When the two reading scores for each question are combined, the correlation between the question scores is .42. This figure represents the degree of relationship between the two essay questions. Since the proportion of shared variance is indicated by the square of the correlation, only 18% of the variance in each question was shared with the other. This suggests that the two questions drew upon essentially different types of writing ability. However, some of the underlying relationship between the two questions was obscured by the unreliability of the grading. An estimate of the true strength of this relationship can be obtained by removing the effect of the unreliability in the grading procedure. When the influence of grading unreliability is extracted from both questions by a procedure known as "correction for attentuation,"[6] the estimated relationship between the questions is .58. The correction for grading unreliability increases the proportion of shared variance to only 34%, which tends to confirm the inference that the two questions represented different domains of content, and sampled different facets of writing ability which are presumably developed in English Composition classes.

[6]D. Magnusson, *Test Theory* (Reading, Mass.: Addison-Wesley, 1967), pp. 147-149.

Correlations Between Parts.

Correlations among the objective, raw essay total, and converted essay scores were computed for the total group. The objective and converted essay scores correlated .53, which is about optimal for an examination of this kind. If the relationship were much lower it would be difficult to justify the compensation which is allowed by combining the scores to form a composite. A substantially higher correlation would mean that each part overlapped the other to the extent that the use of both parts would be redundant.

The correlation of .99 between the raw and converted essay scores indicates a strong linear relationship between the scores. As noted above (Section V, 4), this high degree of linearity means that the two score distributions are very similar in shape, and the linear equating method could have been used without producing serious distortion.

Reliability.

1. **Reliability of the Objective Part.** The reliability of the CLEP "Analysis and Interpretation of Literature" examination is .88 as reported in the CLEP Score Interpretation Guide.[7]

2. **Reliability of the Essay Part.** Because essay tests generally consist of only one or at most several items, it is not possible to compute reliability by split-half or internal consistency methods. The only alternative is "form-to-form" reliability, which entails giving two equivalent forms of a test to the same group of students. However, the preparation of an equivalent form for a one-or-two question test is not feasible, because one or two questions do not constitute an adequate sample from the domain of content. In lieu of a formal method of computing reliability, a method of estimation is available. Upper and lower limits of reliability can be established, with the certainty that the actual value lies *somewhere between* the two limits.

3. **The Upper Limit.** The correlation between grades on the first and second readings of each essay question represents the reliability of the grading procedure. Since unreliability due to grading is only one of several sources of error variance, the value for grading reliability will always be higher than the overall reliability value which includes all sources of error variance. Thus, grading reliability can be used as the upper limit of overall or actual reliability. For Question 1, the correlation between grades on both readings is .51, and for Question 2 the correlation is .61. These values are the reliabilities of a *single* reading. That is, the reliability of each reading of Question 1 is .51. Adding the reading scores to obtain a question score produces an increase in reliability, which is estimated by the Spearman-Brown formula.[8] The estimated reliability for Question 1 is .68, and that for Question 2 is .76.

   Because the essay total score is the sum of the question scores, the reading reliability for the total is higher than that for either of the questions. The reading reliability of the essay total score is most appropriately estimated by the formula for computing the reliability of a linear composite, since the two questions (as established above) are not alternative forms representing the same domain of content.[9] The reading reliability of the essay total score is estimated to be .81, and this value is taken as the upper limit of reliability for the total essay test.

4. **The Lower Limit.** The obtained correlation between Question 1 and Question 2 is used to estimate the lower limit of reliability. If each question were treated as a one-question test, the actual reliability per question could be obtained by giving an equivalent form of a question and correlating scores on both forms. Unreliability due both to grading and to fluctuations in performance would be reflected in such a "form-to-form" coefficient. It has already been demonstrated that Question 1 and Question 2 are not equivalent in terms of content. To the extent that the two questions are not equivalent, the correlation between them will underestimate the actual reliability of the question scores.

[7] , *Score Interpretation Guide,* College-Level Examination Program (Princeton: College Entrance Examination Board, 1967), p. 31.

[8] Magnusson, *op. cit.*, p. 74.

[9] J. Nunnally, *Psychometric Theory* (New York: McGraw-Hill, 1967), pp. 226-230.

The correlation between the scores on Question 1 and Question 2 is .42. Again, because the essay total score includes both questions, the reliability of the total will be higher, and is estimated by the linear composite method to be .59. The actual reliability of the essay test therefore lies between the lower limit of .59 and the upper limit of .81.

5. **Reliability of the Composite Scores.** The estimated reliability of the composite scores, based on reliability coefficients of .88 for the objective part and between .59 and .81 for the essay part, is between .83 and .90.

6. **Reliability in Terms of Reading Score Discrepancies.** Another way of expressing the reliability of the essay scores is to indicate the frequency and magnitude of the discrepancies between the two reading scores for each question. Inasmuch as each of the 2,218 candidates wrote two essays and each essay was read twice, there were 2,218 pairs of scores for each of the two questions. And, as the scores of each reading ranged from 0 to 6, the discrepancy between the two scores in each pair could range from 0 to 6. However, differences of more than 2 points were always resolved by a third reading. Two-point differences were allowed to stand, unless one reading was pass and the other was fail; that is, unless the scores were 5-3 or 4-2. In such cases, a third reading resolved the difference by taking the middle value for both scores or by assigning one of the values for both scores.

   For Question 1, 52.0% of the score pairs had no discrepancy. A discrepancy of one point occurred in 46.4% of the pairs, and a discrepancy of two points occurred in only 1.6% of the cases. For Question 2, the figures were 52.1% for no difference, 45.9% for a one-point difference, and 2.0% for a two-point difference. For the total of 4,436 pairs of reading scores, the percentages were as follows: no difference, 52.1%; one-point difference, 46.1%; two-point difference, 1.8%. In summary, the two reading scores were either identical or differed by only one point in 98.2% of the cases, and two-point differences occurred in the remaining 1.8% of the cases.

Validity.

1. **Validity of the Objective Part.** Course grades of other indicators of classroom achievement can be correlated with test scores to obtain measure of concurrent validity. The CLEP *Score Interpretation Guide* reports correlations between scores on the "Analysis and Interpretation of Literature" examination and course grades before and after the final examination; the values are .42 and .54, respectively. As noted in the CLEP *Guide*,[10] these correlations should be interpreted with caution because of lack of uniformity in the procedures by which the data were obtained.

2. **Validity of the Essay Part.** Course grades earned by students in the CSUC norms samples were correlated with essay total scores. For the combined Fullerton samples, the essay/grade correlation was .40 and for the combined Long Beach samples the value was only .26. Inasmuch as class grades may be affected by factors unrelated to competence, such as absences or incomplete assignments, the instructors were also asked to rank their students on a six-point scale of general ability in the course. It was anticipated that the ranks would more accurately reflect the students' abilities than would the course grades, resulting in a higher correlation with the essay scores. This expectation, however, was only partially realized. The correlations between essay total scores and ranks were .32 for Fullerton and .50 for Long Beach. Thus the use of ranks rather than grades as a criterion resulted in a moderately increased correlation for the Long Beach sample, but produced a slightly diminished value for the Fullerton sample. All the values cited in this section reflect unreliability in grading due to the various standards used by different instructors, and should be interpreted only as gross approximations, if not as underestimates, of concurrent validity.

[10] *op. cit.*, pp. 33, 34.

TABLE 1

FULLERTON NORMS SAMPLE
COMMUNICATIONS 103

Grade	N	Mean	Standard Deviation	Range
A	10	15.00	2.83	8-18
B	16	12.44	2.78	9-17
C	5	10.20	1.72	8-13
A + B	26	13.42	3.07	8-18
Total	31	12.90	3.12	8-18

TABLE 2

FULLERTON NORMS SAMPLE
COMMUNICATIONS 101 & 102

Grade	N	Mean	Standard Deviation	Range
A	4	17.25	1.64	16-20
B	11	15.18	2.52	11-20
C	6	13.67	1.37	12-16
A + B	15	15.73	2.49	11-20
Total	21	15.14	2.42	11-20

TABLE 3

FULLERTON NORMS SAMPLE
ENGLISH 103 & 105

Grade	N	Mean	Standard Deviation	Range
A	6	13.17	3.24	10-19
B	9	13.33	2.58	10-16
C	5	10.00	2.53	5-12
D	2	10.50	4.50	6-15
A + B	15	13.27	2.86	10-19
A + B + C	20	12.45	3.12	5-19
C + D	7	10.14	3.23	5-15
Total	22	12.27	3.32	5-19

TABLE 4

LONG BEACH NORMS SAMPLE
ENGLISH 184

Grade	N	Mean	Standard Deviation	Range
A	2	20.00	0	
B	4	16.50	2.50	13-20
C	9	13.33	3.27	9-18
A + B	6	17.67	2.63	13-20
Total	15	15.07	3.70	9-20

TABLE 5

LONG BEACH NORMS SAMPLE
ENGLISH 300

Grade	N	Mean	Standard Deviation	Range
B	11	14.18	2.04	11-18
C	6	11.33	1.97	8-14
Total	17	13.18	2.43	8-18

TABLE 6

LONG BEACH NORMS SAMPLE
ENGLISH 310

Grade	N	Mean	Standard Deviation	Range
A	4	17.25	2.58	15-21
B	6	17.17	1.77	14-19
C	11	16.27	2.09	14-21
D	4	16.75	1.09	15-18
A + B	10	17.20	1.99	14-21
A + B + C	21	16.71	2.10	14-21
C + D	15	16.40	1.60	14-21
Total	25	16.72	1.97	14-21

TABLE 7

SCORE CONVERSION TABLE
CSUC ENGLISH EQUIVALENCY TEST

May 1977
Essay Total Score – Converted Score

24 – 80
23 – 76
22 – 72
21 – 69

20 – 66
19 – 63
18 – 59
17 – 55
16 – 51

15 – 47
14 – 43
13 – 40
12 – 37
11 – 34

10 – 32
9 – 30
8 – 28
7 – 27
6 – 26

5 – 25
4 – 24
3 – 23
2 – 22
1 – 21

0 – 20

V

FOLLOW-UP STUDIES

The following charts represent third-year results of what is projected as a five-year study of 1973 English Equivalency Examination participants. Of the 4,072 applicants for credit that year 3,273 (or 81%) enrolled on one of the CSUC campuses in fall 1973. The study seeks to ascertain what, if any, differences may obtain in the performance in college of three groups: those taking the test and passing (to receive 6 semester units of credit); 2) those taking the test and not passing; and 3) all entering students.

Any conclusions to be drawn from these third-year results must be most tentative. It does seem apparent that passing students are performing better than the other groups as measured by GPA; that the passing students are enrolling for more credit units than the average of all students and for about as many credit units as those who took but did not pass the test; that those passing do enroll for English courses, despite their credit by examination, and seem to be taking more English courses than do the test applicants who did not receive credit. Even these findings may need to be altered as data on the students' performance accumulate, however.

Those who wish to follow the progress of this follow-up study will find much of interest in the following tables and the additional tables in future reports. The tables published in the 1975 and 1976 volumes are not reprinted here and should also be consulted. In any event, no final conclusions should be drawn until the data are complete. This report was prepared by the CSUC Division of Institutional Research, April 1977.

TABLE I

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Progress of Freshman Participants
as of Fall 1976 by Sex

	Not Enrolled Fall 1976	Below Normal Progress	Normal Progress (Seniors)	Graduated Students (Terminal)	Graduate Students	Total
Passers – N	376	259	423	13	7	1,078
Percent	34.9	24.0	39.2	1.2	0.6	100.0
Failers – N	789	775	604	2	5	2,175
Percent	36.3	35.6	27.8	0.1	0.2	100.0
Passers by Sex						
Male – N	93	84	142	5	0	324
Percent	28.7	25.9	43.8	1.5	–	100.0
Female– N	283	175	281	8	7	754
Percent	37.5	23.2	37.3	1.1	0.9	100.0
Failers by Sex						
Male – N	366	394	278	2	1	1,041
Percent	35.2	37.8	26.7	0.2	0.1	100.0
Female– N	422	381	326	1	4	1,134
Percent	37.2	33.6	28.7	0.1	0.4	100.0

TCSUC – Institutional Research
April 1977

TABLE II

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Progress of Freshman Participants
as of Fall 1976 by Ethnicity

	Not Enrolled Fall 1976	Below Normal Progress	Normal Progress (Seniors)	Graduated Students (Terminal)	Graduate Students	Total
Passers by Ethnicity						
White/Unknown – N	362	245	397	13	7	1,024
Percent	35.4	23.9	28.8	1.3	0.7	100.0
Minority – N	14	14	26	0	0	54
Percent	25.9	25.9	48.1	–	–	100.0
Failers by Ethnicity						
White/Unknown – N	718	705	543	2	4	1,972
Percent	36.4	35.8	27.5	0.1	0.2	100.0
Minority – N	70	70	61	1	1	203
Percent	34.5	34.5	30.0	0.5	0.5	100.0

TCSUC – Institutional Research
April 1977

41

TABLE III

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Mean Earned SEU and GPA for Participants
Enrolled Fall 1976

	Passed		Failed		Total	
	Avg.	N	Avg.	N	Avg.	N
Transfer SEU*	10.9	619	8.3	1,244	9.2	1,863
Total SEU*	87.8	619	83.1	1,244	84.7	1,863
Transfer GPA**	3.3	211	3.1	423	3.2	634
Total GPA**	3.3	590	2.9	1,190	3.0	1,780

*CSU Hayward and CSU Long Beach not included.

**CSC Bakersfield, CSU Hayward, CSU Long Beach, and CSC San Bernardino not included.

TCSUC – Institutional Research
April 1977

TABLE IV

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Mean Earned SEU and GPA for June 1976 Graduates

	Passed		Failed		Total	
	Avg.	N	Avg.	N	Avg.	N
Transfer SEU	14.3	18	26.0	14	19.4	32
Total SEU	129.0	18	135.6	14	131.9	32
Transfer GPA	3.4	8	3.2	7	3.3	15
Total GPA	3.5	18	3.2	14	3.4	32

TCSUC – Institutional Research
April 1977

TABLE V

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Mean Transferred and Total S.E.U. for
Participants Enrolled Fall 1976 by Fall 1973 Enrolled Campus

	Transfer S.E.U.		Total S.E.U.	
Campus	Avg.	N	Avg.	N
Bakersfield	5.4	30	83.0	30
Chico	15.9	137	89.7	137
Dominguez Hills	3.3	18	79.9	18
Fresno	18.6	65	88.7	65
Fullerton	8.1	185	81.7	185
Hayward	NA	NA	NA	NA
Humboldt	6.0	119	78.4	119
Long Beach	NA	NA	NA	NA
Los Angeles	13.3	40	86.9	40
Northridge	9.0	171	83.4	171
Pomona	5.9	131	89.6	131
Sacramento	7.6	102	78.9	102
San Bernardino	10.6	15	80.0	15
San Diego	11.2	221	84.9	221
San Francisco	10.1	146	76.8	146
San Jose	5.3	208	87.8	208
San Luis Obispo	9.1	231	89.7	231
Sonoma	4.0	28	78.7	28
Stanislaus	11.1	16	91.6	16
All Campuses	9.2	1,863	84.7	1,863

TCSUC – Institutional Research
April 1977

TABLE VI

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Mean Transferred and Total S.E.U for
Participants Enrolled Fall 1976 by Fall 1973 Major

Enrolled Major	Transfer S.E.U. Avg.	Transfer S.E.U. N	Total S.E.U. Avg.	Total S.E.U. N
Agriculture	11.3	82	82.2	82
Architecture	7.2	22	92.7	22
Area Studies	4.2	5	88.8	5
Biological Sciences	8.9	145	84.5	145
Business	8.2	278	85.0	278
Communications	9.4	105	86.5	105
Computer Sciences	12.3	27	86.7	27
Education	9.2	87	84.2	87
Engineering	7.2	95	90.6	95
Fine Arts	9.0	126	85.1	126
Foreign Languages	8.3	22	82.4	22
Health	12.5	103	87.7	103
Home Economics	14.0	48	87.6	48
Letters	8.8	83	82.7	83
Mathematics	5.2	33	88.0	33
Physical Sciences	6.1	60	84.1	60
Psychology	11.3	74	83.8	74
Public Affairs	11.0	75	84.5	75
Social Sciences	8.3	191	83.4	191
Interdisciplinary	10.5	122	86.7	122
Undeclared	6.2	80	71.1	180
All Majors	9.2	1,863	84.7	1,863

TCSUC – Institutional Research
April 1977

TABLE VII

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Mean Transferred and Total GPA for
Participants Enrolled Fall 1976 by Fall 1973 Enrolled Campus

Campus	Transfer GPA		Total GPA	
	Avg.	N	Avg.	N
Bakersfield	NA	NA	3.2	77
Chico	3.1	63	3.1	136
Dominguez Hills	3.0	5	3.1	17
Fresno	3.4	29	3.2	60
Fullerton	3.3	67	3.0	183
Hayward	NA	NA	NA	NA
Humboldt	3.2	33	3.1	110
Long Beach	NA	NA	NA	NA
Los Angeles	3.0	15	3.0	40
Northridge	3.1	66	3.0	168
Pomona	3.0	14	2.9	131
Sacramento	3.3	40	3.1	95
San Bernardino	NA	NA	NA	NA
San Diego	3.1	105	3.0	219
San Francisco	3.4	40	3.1	129
San Jose	3.2	73	3.0	197
San Luis Obispo	3.1	74	3.0	228
Sonoma	2.9	7	3.4	25
Stanislaus	2.9	3	3.3	15
All Campuses	3.2	634	3.0	1,780

TCSUC – Institutional Research
April 1977

TABLE VIII

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Mean Transferred and Total GPA for
Participants Enrolled Fall 1976 by Fall 1973 Major

	Transfer GPA		Total GPA	
Enrolled Major	Avg.	N	Avg.	N
Agriculture	3.1	31	2.9	81
Architecture	2.9	4	2.8	22
Area Studies	4.0	1	3.4	5
Biological Sciences	3.1	56	2.9	143
Business	3.2	99	3.0	267
Communications	3.0	38	3.0	103
Computer Sciences	3.1	10	3.0	27
Education	3.2	27	3.0	81
Engineering	3.1	23	2.9	92
Fine Arts	3.4	36	3.2	117
Foreign Languages	3.8	2	3.3	19
Health	3.2	50	3.1	98
Home Economics	3.0	21	3.0	48
Letters	3.2	28	3.1	77
Mathematics	3.1	8	3.2	31
Physical Sciences	3.1	13	3.0	56
Psychology	3.1	37	3.0	73
Public Affairs	3.3	29	3.1	73
Social Sciences	3.2	60	3.1	180
Interdisciplinary	3.3	42	3.2	116
Undeclared	3.2	19	2.6	71
All Majors	3.2	634	3.0	1,780

TCSUC – Institutional Research
April 1977

TABLE IX

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Changes in Campus and Enrolled Major
Fall 1973 to Fall 1976

	Enrolled at Original Campus	Transferred	Unenrolled	Total
Passers – N	610	89	405	1,104
Percent	55.3	8.0	36.7	100.0
Failers – N	1,212	183	809	2,204
Percent	55.0	8.3	36.7	100.0

	Enrolled in Original Major	Changed Major	Unenrolled	Total
Passers – N	426	273	405	1,104
Percent	38.6	24.7	36.7	100.0
Failers – N	875	520	809	2,204
Percent	39.7	23.6	36.7	100.0

TCSUC – Institutional Research
April 1977

TABLE X

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC ENROLLED PARTICIPANTS

Intercampus Transfer Activity
Fall 1973 to Fall 1976

Campus	Fall 1973 Enrollment	Number Tfrd. In	Number Tfrd. Out*	Net Change	Percent Change	No Longer Enrolled	Fall 76 Enrollment
Bakersfield	55	4	6	-2	-3.6	23	30
Chico	221	26	20	6	2.7	90	137
Dominguez Hills	33	0	3	-3	-8.1	12	18
Fresno	86	13	4	9	10.5	30	65
Fullerton	312	16	21	-5	-1.6	122	185
Hayward	118	10	13	-3	-2.5	47	68
Humboldt	255	7	31	-24	-9.4	112	119
Long Beach	207	27	10	17	8.2	61	163
Los Angeles	68	9	3	6	8.8	34	40
Northridge	272	14	21	-7	-2.6	94	171
Pomona	215	12	19	-7	-3.3	77	131
Sacramento	138	17	8	9	6.5	45	102
San Bernardino	31	4	4	0	0.0	16	15
San Diego	344	28	32	-4	-1.2	119	221
San Francisco	217	28	12	16	7.3	87	146
San Jose	275	26	16	10	3.6	77	208
San Luis Obispo	332	28	26	2	0.6	103	231
Sonoma	96	2	18	-16	-16.7	52	28
Stanislaus	33	1	5	-4	-12.1	13	16
All Campuses	3,308	272	272	0	–	1,214	2,094

*Students no longer enrolled not included.

TCSUC – Institutional Research
April 1977

TABLE XI

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC PARTICIPANTS

Fall 1973 Enrolled Campus by Fall 1976 Enrolled Campus

Fall 1976	BAK	CHI	D H	FRE	FUL	HAY	HUM	L B	L A	NOR	POM	SAC	S B	S D	S F	S J	SLO	SON	STA	Unenrolled	Fall 1973 All Campuses
Fall 1973																					
Bakersfield	26	0	0	4	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	23	55
Chico	0	111	0	1	1	1	0	2	0	0	0	5	0	3	4	1	1	0	1	90	221
Dominguez Hills	0	0	18	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	12	33
Fresno	0	0	0	52	0	0	0	0	0	0	0	0	0	1	0	1	2	0	0	30	86
Fullerton	0	0	0	0	169	1	0	6	2	0	5	0	0	1	1	1	4	0	0	122	312
Hayward	0	3	0	0	0	58	0	0	0	0	0	2	0	1	1	4	2	0	0	47	118
Humboldt	1	7	0	1	2	3	112	3	0	0	0	2	0	2	3	4	2	1	0	112	255
Long Beach	0	1	0	0	1	0	0	136	0	1	1	1	1	2	0	0	2	0	0	61	207
Los Angeles	0	0	0	0	0	0	0	1	31	0	0	0	1	0	0	1	0	0	0	34	68
Northridge	0	0	0	0	1	0	0	1	2	157	0	1	1	10	1	2	2	0	0	94	272
Pomona	1	1	0	0	4	0	1	1	4	0	119	0	0	2	0	0	5	0	0	77	215
Sacramento	1	1	0	1	0	0	1	0	0	1	0	85	0	0	1	1	1	0	0	45	138
San Bernardino	0	1	0	0	0	0	0	0	0	0	2	1	11	0	0	0	0	0	0	16	31
San Diego	0	4	0	0	4	1	2	8	0	5	1	0	0	193	1	3	3	0	0	119	344
San Francisco	0	3	0	1	1	0	1	0	1	1	0	0	0	1	118	3	0	0	0	87	217
San Jose	0	0	0	0	1	4	0	2	0	0	0	2	0	2	3	182	2	0	0	77	275
San Luis Obispo	1	3	0	4	1	0	0	2	0	3	3	2	1	2	1	2	203	1	0	103	332
Sonoma	0	1	0	0	0	0	1	0	0	3	0	1	0	0	10	2	0	26	0	52	96
Stanislaus	0	1	0	1	0	0	1	0	0	0	0	0	0	1	0	1	0	0	15	13	33
All Campuses	30	137	18	65	185	68	119	163	40	171	131	102	15	221	146	208	231	28	16	1,214	3,308

TABLE XII

FALL 1973 ENROLLED ENGLISH EQUIVALENCY PARTICIPANTS
CHANGES IN ENROLLMENT BY MAJOR
FALL 1973 TO FALL 1976

Major	Fall 73 Enrollment	Changes to Major	Changes From Major*	Net Change	Percent Change	No Longer Enrolled	Fall 76 Enrollment
Agriculture	141	32	36	-4	-2.8	55	82
Architecture	32	8	12	-4	-12.5	6	22
Area Studies	4	5	2	3	75.0	2	5
Bio. Sciences	348	56	104	-48	-13.7	130	170
Business	284	178	60	118	41.5	82	320
Communications	112	53	23	30	26.8	31	111
Computer Science	17	17	2	15	88.2	5	27
Education	108	60	40	20	18.5	32	96
Engineering	140	34	29	5	3.6	43	102
Fine Arts	231	52	45	7	3.0	95	143
Foreign Languages	44	18	16	2	4.5	18	28
Health	129	64	28	36	27.9	40	125
Home Economics	55	32	12	20	36.4	20	55
Letters	165	58	70	-12	-7.3	60	93
Mathematics	91	16	37	-21	-23.1	32	38
Physical Sciences	96	41	32	9	9.4	36	69
Psychology	117	61	47	14	12.0	42	89
Public Affairs	43	69	9	60	139.5	16	87
Social Sciences	280	126	71	55	19.6	121	214
Interdisciplinary	72	101	20	81	112.5	21	132
Undecided	799	39	425	-386	-48.3	327	86
All Majors	3,308	1,120	1,120	0	–	1,214	2,094

*Students no longer enrolled not included.

TCSUC – Institutional Research
April 1977

52

TABLE XIII

ENGLISH EQUIVALENCY EXAMINATION
1973 CSUC PARTICIPANTS

Fall 1973 Major by Fall 1976 Major

Fall 1976	Agri.	Arch.	Area Stds.	Bio. Sci.	Busi-ness	Comms.	Comp. Sci.	Educ.	Engr.	Fine Arts	For. Langs.	Health	Home Econ.	Letters	Math.	Phys. Sci.	Psych.	Pub. Aff.	Soc. Sci.	Inter-disc.	Undec.	Unen-rolled	Fall 1973 All Majors
Fall 1973																							
Agriculture	50	1	0	6	4	1	0	1	1	0	0	2	1	1	1	0	1	2	1	0	13	55	141
Architecture	3	14	0	1	0	2	0	1	2	0	0	0	0	0	2	0	0	0	0	0	1	6	32
Area Studies	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	2	4
Bio. Sciences	7	0	0	114	16	4	2	5	5	2	2	8	2	6	1	10	9	4	13	7	1	130	348
Business	5	0	0	3	142	3	1	1	1	5	3	2	2	4	0	3	2	4	15	3	3	82	284
Communications	1	0	1	2	3	58	1	1	0	1	0	0	0	3	0	1	2	1	5	1	2	31	112
Computer Science	0	0	0	0	0	0	10	0	1	0	0	0	0	0	0	0	0	0	1	0	0	5	17
Education	0	0	0	1	3	0	1	36	0	1	0	8	2	2	0	1	4	3	4	9	1	32	108
Engineering	2	4	0	2	7	0	0	2	68	0	0	0	0	0	1	1	0	0	2	0	8	43	140
Fine Arts	0	0	0	0	4	4	1	3	0	91	1	0	1	3	1	1	3	7	7	6	3	95	231
Foreign Languages	0	0	0	0	4	0	0	0	0	0	10	1	1	2	0	2	1	0	1	3	1	18	44
Health	0	0	0	3	4	2	0	3	0	2	1	61	2	1	0	0	3	3	1	3	0	40	129
Home Economics	1	0	0	0	3	1	0	2	0	0	0	2	23	0	0	1	0	1	0	1	0	20	55
Letters	0	0	0	1	3	7	0	5	0	10	6	4	1	35	0	1	3	3	6	18	2	60	165
Mathematics	1	1	0	2	16	0	3	0	1	3	0	0	2	1	22	2	1	1	1	2	0	32	91
Physical Sciences	3	0	0	5	6	1	2	3	4	0	0	2	0	1	2	28	1	0	2	0	0	36	96
Psychology	0	0	0	2	6	1	0	5	0	4	0	2	0	5	0	0	28	8	7	7	0	42	117
Public Affairs	0	0	0	0	1	0	0	2	0	0	1	1	0	1	0	0	0	18	2	1	0	16	43
Social Sciences	1	1	1	3	16	9	0	5	2	2	0	2	2	2	1	3	3	7	88	9	2	121	280
Interdisciplinary	0	0	0	0	3	2	1	0	2	2	0	2	0	0	0	0	0	1	5	31	2	21	72
Undecided	8	1	3	27	79	16	5	21	15	20	4	28	16	26	7	15	28	24	52	30	47	327	799
All Majors	82	22	5	170	320	111	27	96	102	143	28	125	55	93	38	69	89	87	214	132	86	1,214	3,308

TCSUC – Institutional Research
April 1977

APPENDIX A

ENGLISH EQUIVALENCY EXAMINATION 1976-77

BUDGET SUMMARY

Start-Up Costs:

Personal Services	$22,466
Operating Expenses	13,500
TOTAL	$35,966

Note: The above includes administrative time, secretarial and student assistant help, printing (brochure and essay test questions), mailing, and travel (including that of various faculty committees).

Reading Costs:

Readers and Professional Team	$17,000
Aides	1,865
Operating Costs	9,268
TOTAL	$28,133

Note: The above operating expenses include transportation and per diem.

Follow-Up Costs:

Summer Administration and Preparation of Final Report	$ 6,078
Summer Clerical Assistance	550
Summer Operating Expenses	250
TOTAL	$ 6,878

Assistance underwritten by The California State University and Colleges campuses of San Jose, Pomona, and San Bernardino; and by the Office of the Chancellor.

Office Space, San Jose and San Bernardino
Telephone Usage, San Jose
Computer Time, Pomona
Key Punch Operator, Pomona
Statistician, Pomona
Computer Time for Follow-Up Study, Chancellor's Office
Development of Follow-Up Study, Chancellor's Office

APPENDICES

State of California

Memorandum

California State College, San Bernardino
5500 State College Parkway, San Bernardino, California 92407

To: David H. Provost

Date: October 24, 1977

From: Edward M. White
English Equivalency Examination

Subject: Increasing student fees for the English Equivalency Examination

The Chancellor's recent inquiry about the proportion of student support for the English Equivalency Examination has led us to a careful review of the EEE budget and program. Our conclusion is that the CSUC is getting a very substantial return on its investment in this program, and that to increase the present $20 student fee would be a most unfortunate economy.

We have, in fact, for some years been seeking ways to reduce that fee, particularly for less affluent students, in order to make this opportunity for credit by examination more widely accessible. As the test, and its high academic standards, have become familiar to students, the number attempting it has decreased (from 1973: 4,072 to 1977: 2,218). If we were to ask students to pay more than the existing $20 test fee, I fear the decrease in applicants might jeopardize the viability of the program.

If students were to absorb the cost of the essay reading, and we were to assume a minimal 10% decrease in enrollment due to the fee increase, approximately 2,000 students would need to cover approximately $30,000. The fee would jump to $35 for the test. Such a charge is, on the face of it, unrealistic. Even though it is clear that $35 is not exorbitant for six semester units of credit, it is also clear that we would be pricing the test beyond a reasonable risk level for most high school students. They would be likely to seek credit by some other means. Unfortunately, the only other generally acceptable vehicle for English composition credit in the CSUC is the CEEB Advanced Placement Program, which is not available at many California high schools.

I think it important that we keep in mind the high professional respectability of the EEE, and its profound influence upon English testing nationwide. It is our program, so I have been informed by key ETS people, that has demonstrated the viability of student writing on writing tests, and led to a return to writing on some national test programs. That is value for money expended. Again, I think we should value highly the enhanced professional development that the EEE has accomplished

57

in our own English departments. Stimulated by the test, our English departments have accomplished substantial innovations in their programs. Because EEE participation is so widespread, particularly among departmental leaders (in 1977, 22% of essay readers were department chairs or composition directors), our procedures and practices immediately influence English instructors throughout the CSUC.

It is no accident that the major effort to raise CSUC standards this last year, through the development and use of the English Placement Test, was sponsored and supported throughout by our English departments. Their test expertness and experience have led to this new and humane effort to upgrade the CSUC degree, and, in effect, upgrade all California secondary education.

Finally, of course, the example of the EEE has spread to other disciplines. Most recently, the mathematics-science equivalency test has opened new opportunities for ambitious and able students, and the model of the EEE was crucial for this new program. The College Board has recognized this radiating effect of the EEE by reprinting a theoretical appendix from an EEE report (Comparison and Contrast, 1975, pp. 101-111) in The College Board Review and then distributing it at a series of regional meetings for faculty interested in advanced placement programs.

The CSUC has received very substantial value for its investment in the EEE, and I hope that the test program receives continued support from the CSUC. I can think of no route more likely to destroy the EEE than to raise student fees. I continue to hope that ways can be found to decrease the fee, for the major potential of the program has still not been realized.

APPENDIX B

THE 1977 ENGLISH EQUIVALENCY RELIABILITY STUDY

In determining reliability we are concerned with the stability of ratings or scores over repeated instances of measurement. More precisely, we are concerned that a given level of performance will be assigned the same or essentially the same score on different occasions. Most studies of the reliability of essays deal only with the reliability of the scoring process, and consist of the rescoring of the same set of papers by the same team of readers, or in some cases, by a different team of readers. In situations where the same team of readers is used for the second reading, it is desirable that the papers be so distributed that no reader rescores a paper he has previously marked. Finlayson[1] has noted that when the same team of readers awards both sets of marks, the reliability coefficient will be spuriously high because the idiosyncratic tendencies of the readers will affect the second set of readings in the same manner as the first set was affected. If it were common practice to employ the same group of readers for repeated administrations of a test, the determination of reliability by using the same readers would not be potentially misleading, inasmuch as it would correspond to actual educational practice. It is not realistic, however, to expect that the same team of readers could be used repeatedly in a testing program. Thus a reliability coefficient based on two sets of marks by the same team of readers would be inflated in comparison to the coefficient between scores obtained from two separate test administrations.

In essay testing, certain constraints are always encountered with respect not only to the reemployment of the same group of readers but also with respect to the repetition of the same topic. As noted above, practical considerations ordinarily dictate that the composition of the reading team will change from reading to reading, and for security purposes it is usually not advisable to use the same topic repeatedly. Given these constraints in actual practice, it can be argued that a reliability study should approximate the conditions prevailing in an ongoing testing program as closely as possible. Otherwise the obtained coefficient and any inferences about the stability of scores must be made in a hypothetical context which may have little practical meaning.

On the basis of the foregoing considerations, Finlayson[2] has suggested that any overall measure of essay reliability should be a test re-test correlation. This would require that two separate essay performances by each student be scored, with each student writing on two topics very similar in content but different in details. Unfortunately, such ideal conditions are not often approximated in reliability studies. Vernon and Millican,[3] however, conducted a comprehensive reliability study in which they obtained average coefficients for the following situations: 1) same essay, correlation between two markers (.509); 2) same markers, correlation between different topics (.366); 3) different topics, correlation between different markers (.224).

The topics in the Vernon and Millican study were in no sense designed to be parallel, and called for rather different types of expository writing. Therefore, the correlations between different topics (for either the same markers or different markers) are lower than would be expected if the questions had been parallel in content, as would be desirable in a test re-test study. The authors concluded that they found no consistent English ability when different essay topics were scored by different readers. However, they did achieve rather high reliability (.83) when the score for each writer was the sum of two readings each of seven different essay topics. This study illustrates the great range of approaches that can be taken in determining reliability. On the one hand, the correlation of .509 between two markers scoring the same essay is slightly lower (but obtained by a roughly comparable method) than the inter-reader correlations obtained in the English Equivalency reading. On the other hand, the value of .83 obtained by summing two readings over seven different papers represents a sheer quantity of reading that would not be practical in an actual testing situation.

[1] D. S. Finlayson, "The Reliability of the Marking of Essays," *British Journal of Educational Psychology*, 1951, *21*, 126-134.

[2] *Ibid.*

[3] P.E. Vernon, and G. D. Millican, "A Further Study of the Reliability of English Essays," *British Journal of Statistical Psychology*, 1954, *VII*, 65-74.

Often it is not possible to conducted an essay testing program in such a manner that reliability can be calculated on the basis of parallel questions and different teams of readers. In the case of the English Equivalency Exam, it is only feasible to calculate the reading or scoring reliability. If the second reading of the essay papers takes place immediately following the first reading, and the same team of readers grades the papers both times, the highest possible value for reading reliability will be obtained. Both readings will be conducted by the same readers (although no reader will read the same paper twice), according to the same standards, at the same time of day, at the same point in the reading, and in the same psychological milieu. Such has been the case with the EEE essay reading. The kind of practical questions that should be asked about reliability would go something like this – "Would a student who obtained a barely passing score (14) on the essay in 1976 pass the test if he took it in 1977?" Such a question would presuppose, of course, that there would be no intellectual or educational growth on the part of the student between 1976 and 1977. A partial answer to this question could be obtained if we administered the 1976 questions to a group of 1977 examinees under actual testing conditions. Ideally, a random half of the examinees would write the 1976 questions first, and the other half would write the 1977 questions first. However, the question posed above could not be answered with certainty because two important factors would not remain constant: 1) the composition of the reading team would change from 1976 to 1977; and 2) the cognitive and affective atmosphere of the 1977 reading would be somewhat different than that which prevailed in 1976.

In an attempt to determine the effect of time lapse and a substantial change in the composition of the reading team on scoring reliability, the 1977 English Equivalency reading team rescored a sample of 1976 papers. At the conclusion of the 1977 reading, 699 1976 essay papers were distributed to the readers for rescoring. The 1976 grades were covered by a new scoring sheet which was stapled to the face of each essay booklet. Masking tabs were used to cover the score assigned by the first reader, so that the score could not be consulted by the second reader. After a brief training session to reorient the readers to the 1976 questions, the rescoring of the 1976 papers proceeded according to exactly the same procedures which governed the scoring of the 1977 papers.

Several important factors, in addition to the time lapse and changes in the reading team, caused this rescoring procedure to produce lower correlations than the estimates of grading reliability published in 1976. It would no doubt be helpful at this point to list all of the major factors which contributed to the lowered reliability figures. First, the composition of the reading team was altered by the inclusion of 50% new readers. The 1977 team of 36 readers included 18 readers who had not read in 1976; of these 18 new readers, 11 had never participated in an EEE reading before, and 7 had read prior to 1976. Second, the reading took place a year later, and the judgments of the readers may have been slightly shaded by the experiences of the intervening year. Third, all of the rescoring was done at the very end of the 1977 reading, and there is evidence that scoring standards tend to shift during the course of a reading (Godshalk *et al.*, 1966).[4] Although a shift in standards would probably not affect the correlation, it would affect the distribution of discrepancy scores (to be discussed later). Finally, there were undoubtedly subtle changes in the psychological atmosphere of the reading which might have produced measurable changes in the scoring.

Each year, in the regular EEE reliability computations, the scores for the first and second readings of each question are correlated. Because both reading scores are added together to obtain a question score, the coefficient obtained by correlating the first and second readings is expanded by the Spearman-Brown formula so that it represents the total score for each question. In 1976, the correlation between the first and second readings of Question 1 was .58, which was expanded by the Spearman-Brown formula to .73 for the sum of both readings. When Question 1 for 1976 was rescored in 1977, the correlation between the total scores for 1976 and 1977 was .63, somewhat lower than the Spearman-Brown estimate of .73. In the context given, the Spearman-Brown estimate represents the highest possible figure for scoring reliability, with no attenuating factors present. The correlation of .63 obtained by rescoring the 1976 papers represents a more pragmatic or educationally realistic value for the scoring reliability of Question 1 in 1976. (It should be noted that changes in the content of Question 1 from year to year will affect the reliability, although an attempt is made to present the writers with essentially the same demands each year. Thus the figures given apply only to 1976.) The Spearman-Brown value for Question 2 in 1976 was .83, and the value

[4]F. I. Godshalk, F. Swineford, and W. E. Coffman, *The Measurement of Writing Ability* (New York: College Entrance Examination Board), 1966.

obtained by rescoring the papers in 1977 was .76. The estimated reliability for the total essay in 1976 (Question 1 + Question 2) was .85. The total essay reliability obtained by rescoring the papers in 1977 was .78, again slightly lower due to the attenuating factors discussed above.

A caveat should be issued at this point. The interpretation of a reliability coefficient is contingent upon the conditions under which it was derived, and this is particularly so for the reliability of essay readings. Any comparison of reliability values for different essay readings cannot avoid a degree of uncertainty, even when all conditions which could affect the reliability are carefully analyzed. Vernon and Millican[5] list for factors that can influence the reliability coefficients of essays. First, unskilled writers will tend to be graded more for obvious mechanical flaws than for more subjective qualities such as content and style. Thus scores for less adequate writers will tend to be more reliable. Second, highly experienced graders will produce more reliable ratings than will less experienced graders. Third, longer essays will be more reliably read than shorter ones. Fourth, reliability will vary directly with the range of ability among the writers. Vernon and Millican also suggest that the choice of essay topics may affect reliability.

Inasmuch as a coefficient of correlation does not adequately convey the educational significance of reliability, a less technical and more tangible measure of score reliability is needed. One practical means of expressing the stability of scores is to determine the difference or discrepancy between the first and second reading scores for each subject, and then produce a frequency distribution of discrepancy scores. When the first and second reading scores are identical there is no discrepancy. When the scores differ by one scale point, a discrepancy of one exists, and so on. In addition to the mean and standard deviation of the discrepancy distribution, the percentage of scores falling in each discrepancy point category can be calculated. The following tables give the discrepancy distributions for the 1977 rescoring of the 1976 English Equivalency essays. Table 1 gives the distribution for Question 1, Table 2 represents Question 2, and Table 3 gives the total essay score discrepancy distribution. Because each question score is the sum of two readings on a 6-point scale, scores for Questions 1 and 2 fall on a 12-point scale. The total score represents the sum of the scores on Questions 1 and 2, and is therefore expressed on a 24-point scale.

Reliability data expressed in this form enable us to make some concrete observations about the stability of scores. For instance, when Question 1 was rescored in 1977, 31.3% of the scores were the same as in 1976. The 1977 scores differed by only one point in 45.4% of the cases, and by no more than one point in 76.7% (31.3 + 45.4) of the cases. When the 2-point discrepancies are included (17.9%), it is apparent that in 94.6% of the cases the discrepancy was two points or less on the 12-point scale. The discrepancy figures for Question 2 are almost identical to those of Question 1. The 1977 total scores, on a 24-point scale, were identical to the 1976 scores in 20.7% of the cases; 58% of the scores varied by one point or less, 82.7% by two points or less, and 92% by three points or less. The average discrepancy score for both Question 1 and Question 2 was 1.00, and the average discrepancy for the total score was 1.50. The practical significance of these data can be shown by some simple observations. In general, the probability is .58 that students who received a particular total essay score in 1976 would score within one point of that score in 1977; the probability is .827 that students will receive a 1977 score within two points of their 1976 score. Because the percentage of scores in each discrepancy category is an average based on the whole distribution, these probabilities will vary slightly for any given score.

An exact determination of the amount of discrepancy involved in the rescoring of a given 1976 score can be made by consulting Table 4, which presents a scatter-plot of the 1976 and 1977 data. For instance, by locating a 1976 score of 14 on the left margin, and noting the number of different 1977 scores along the corresponding row, we can determine that there were 21 + 20 + 13 = 54 discrepancies of one point or less. These discrepancies represent 54/86 or 62.8% of the data for 1976 scores of 14. Proceeding in like manner, 36/86 or 41.9% of those students who scored 14 in 1976 would have scored less than 14 in 1977. The EEE statistical report for 1977 presents discrepancy data for differences between the first and second readings of Question 1 and Question 2. However, discrepancy data for each question and for the essay total score are not available with the reliability model presented in the statistical report, as such data would require the rescoring of each question and would double the length of the reading.

[5]Vernon and Millican, *op. cit.*

It should be kept in mind that the reliability data discussed above were obtained by rescoring a set of papers after a lapse of one year, and that the composition of the reading team had been substantially altered. If our concern is only with the reliability of the scores at a given point in time with all factors held as constant as possible (such as an immediate rescoring of all papers), the higher estimates of reading reliability published in the statistical report are the appropriate figures. In brief, the interpretation of any reliability coefficient must be restricted to a hypothetical situation exactly parallel to the context in which the coefficient was derived.

TABLE 1

QUESTION 1
DISCREPANCY DISTRIBUTION

Discrepancy	Frequency	Percentage
0	219	31.3
1	317	45.4
2	125	17.9
3	25	3.6
4	11	1.6
5	2	.3

Mean = .996
Standard deviation = .907

TABLE 2

QUESTION 2
DISCREPANCY DISTRIBUTION

Discrepancy	Frequency	Percentage
0	224	32.0
1	295	42.4
2	144	20.6
3	26	3.7
4	7	1.0
5	3	.4

Mean = 1.007
Standard deviation = .914

TABLE 3

TOTAL ESSAY SCORE
DISCREPANCY DISTRIBUTION

Discrepancy	Frequency	Percentage
0	145	20.7
1	261	37.3
2	173	24.7
3	65	9.3
4	33	4.7
5	18	2.6
6	4	.6

Mean = 1.499
Standard deviation = 1.249

TABLE 4

SCATTERPLOT OF EEE SCORES
RELIABILITY SAMPLES 1976 & 1977

Distribution of 1977 Scores

Distribution of 1976 Scores	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	N	$\bar{X}$
4						1																1	9.0
5	1		4																			5	5.6
6		2	1	1	2	2																8	7.1
7				1	1	4	1	1	1													9	9.3
8				2	2	6	5	1	2	1												19	9.6
9				2	3	10	9	11	4	3	1	1										44	10.3
10				1	5	9	6	15	10	1	1	2	1									51	10.7
11					1	5	10	22	26	9	8	2	1									84	11.7
12						3	12	16	22	18	14	4	2	1	1							93	12.3
13					1	1	2	8	15	20	26	13	4	2								92	13.3
14							2	3	11	20	21	13	9	4	3							86	14.0
15							2	2	6	12	13	17	4	6	2							64	14.2
16								2	3	7	11	13	12	7	4		1					60	15.0
17										1	3	8	13	4	4							33	15.9
18									1	1	5	3	4		2	1		1				18	15.6
19											2	1	1	3	3	2	1					13	17.1
20												1		1	3	4	1					10	18.2
21													2		1	1		2	1			7	19.0
22																	1					1	20.0
23																					1	1	24.0
24																							
N	1	2	5	7	15	41	49	81	101	93	105	78	53	28	23	8	4	3	1		1	699	
$\bar{X}$	5.0	6.0	5.2	8.1	9.1	9.2	10.7	11.2	12.0	13.2	13.9	14.6	15.6	15.8	17.0	19.6	19.3	20.0	21.0		23.0		

1976 DATA

Overall mean = 12.991

Overall standard deviation = 3.087

1977 DATA

Overall mean = 13.017

Overall standard deviation = 2.774

APPENDIX C

SON OF EEE: THE ENGLISH PLACEMENT TEST

When the CSUC Board of Trustees voted in May of 1976 to require a "proficiency/diagnostic test" of all lower-division students, it was responding to a problem within a context. The problem was, of course, the well-documented decline in writing skills of entering college freshmen, supported by widespread CSUC faculty discontent with the writing of undergraduates in all classes. (Surveys taken by a systemwide task force the preceding year documented this discontent.) The context was in part established by the English Equivalency Examination, with its three-year history of demonstrated credibility and systemwide participation.

It was clear from the outset that the EEE, in its present form, was not an appropriate placement test for entering students. A placement test seeks to discover student readiness to *enter* a course of study; an equivalency test measures performance according to norms set by those students *concluding* a course of study. Furthermore, the content of a placement test depends upon the definition of skills and learning required of students who are to profit from a curriculum; the content of an equivalency test depends upon the definition of a curriculum and its goals. The EEE in its present form was developed out of a lengthy process of consultation with the nineteen CSUC freshmen English coordinators, and its goal is to reflect the goals of freshman English instruction in general in the system. (See *Comparison and Contrast,* 1973, for the description of the development of the EEE.) In particular, it was not to be expected that a significant percentage of entering freshmen could reach acceptable score levels on either the expository-analytic essay (Question 2), or the CLEP Analysis and Interpretation of Literature multiple-choice test.

However, by calling for a "proficiency/diagnostic test," the Trustees voiced a commonly held expectation that a new test might be devised to meet both the placement needs of the new program and the equivalency needs of the existing EEE. Thus one of the first questions the EPT Development Committee faced had to do with this expectation: Could the new test reliably determine placement at the lower score levels and equivalency at the upper score levels? The question could be well handled, since all but one of the new test committee members had substantial experience working with the EEE, as well as with classes of marginally qualified freshmen: the committee came together with direct experience of the requirements of both goals.[1]

After substantial careful discussion, the EPT development committee regretfully concluded that it would not be possible at this point to ask a single test to accomplish both English placement and freshman English equivalency. While the EEE appropriately reflects requirements and performance levels of college freshman English courses, a placement test needs to measure a carefully defined and delimited set of skills. Furthermore, a placement test needs to be rather easy, so that different kinds of poorly prepared students may be distinguished and identified for special assistance, and so that diagnostic information about the basic skills of such students may be obtained; an equivalency examination needs to be rather difficult, so that fine distinctions may be made among the best prepared students, and so that college credit may be awarded only to the truly deserving.

Thus the EPT development committee requested that the new test be called a "placement test," rather than a "proficiency test," since it was designed specifically to assist in placement of students with marginal abilities in reading and writing. Those scoring at the upper ranges should consider themselves ready for college composition, not necessarily "proficient."

The EPT consists of four parts:

I. Reading
II. Sentence Construction
III. Logic and Organization
IV. Essay

[1]The EPT Development Committee for 1976-78: Professor Rex Burbank, Chair, San Jose; Professor Charles Moore, CSU, Sacramento; Professor Jill Cohn, CSC, Bakersfield; Professor Rosentene Purnell, CSC, Stanislaus; Professor David Rankin, CSU, Dominguez Hills; and Professor William Robinson, San Francisco State University.

The descriptive brochure for students points out carefully the difference between the placement test and other tests; such as the EEE:

> Because the test is meant, first of all, to determine whether you can enter college-level classes without a handicap in reading or writing sufficiently severe to prevent you from succeeding, it is not as difficult as some other English tests. That is, this test is designed to find out whether you have enough skill in reading and in written communication to undertake college-level work and not whether you have learned all that might be expected of the student who will excel in college courses.

EPT score reports for each student include four sub-scores, one for each portion of the test; and two combined scores, giving Composition and Total scores. The test was first administered in August, 1977, to over 12,000 students; test results were made available to the students and to their campuses in time for Fall, 1977 advising and enrollment. The 1977-78 academic year will include extensive evaluation of this new placement test and its uses. By 1978-79, all entering lower division students (about 50,000 a year) will be required to take the test, and all nineteen campuses will have instituted programs, within their funded capacity, responding to the needs of students made manifest by the test results.

Information about the EPT may be obtained from Dean Anthony Moye, at the CSUC Chancellor's Office, 400 Golden Shore, Long Beach, California 90802, or from the Educational Testing Service, which is handling test operations for the CSUC.

The following tables present data describing the group of students who participated in the 1977 administration of the English Equivalency Examination. Tables are included which detail performance data and age characteristics of four groups of participants: (1) those who passed for unit credit, (2) all those who were not awarded unit credit, (3) those who did not earn unit credit but who scored highly enough to meet the English Placement Test requirement, and (4) those who received neither unit credit nor EPT credit. The third and fourth groups are subgroups of the second.

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES

English Equivalency Examination – Spring 1977

Total Applicants by Age

Age	Non-Passers			Placement Passers			Credit Passers			Total		
	N	Rel. %	Cum. %	N	Rel. %	Cum. %	N	Rel. %	Cum. %	N	Rel. %	Cum. %
16 and Under	1	0.4	0.4	4	0.4	0.4	5	0.6	0.6	10	0.5	0.5
17	12	4.9	5.3	73	6.7	7.1	65	7.4	7.9	150	6.8	7.2
18	162	66.1	71.7	815	74.8	82.1	606	68.6	76.7	1,583	71.6	78.8
19	40	16.3	88.1	112	10.3	92.4	79	8.9	85.7	231	10.4	89.3
20	3	1.2	89.3	11	1.0	93.5	14	1.6	87.3	28	1.3	90.5
21	2	0.8	90.2	13	1.2	94.7	12	1.4	88.6	27	1.2	91.8
22	2	0.8	91.0	3	0.3	94.9	16	1.8	90.5	21	0.9	92.7
23	2	0.8	91.8	3	0.3	95.2	14	1.6	92.1	19	0.9	93.6
24	1	0.4	92.2	4	0.4	95.6	8	0.9	93.0	13	0.6	94.2
25-29	2	0.8	93.0	12	1.1	96.7	18	2.0	95.0	32	1.4	95.6
30-34	3	1.2	94.3	2	0.2	96.9	11	1.2	96.3	16	0.7	96.3
35 and Over	14	5.7	100.0	34	3.1	100.0	33	3.7	100.0	81	3.7	100.0
All Ages	244*	100.0	–	1,086*	100.0	–	881*	100.0	–	2,211*	100.0	–

*Age data unobtainable from 7 applicants

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES

English Equivalency Examination – Spring 1977

Male Applicants by Age

Age	Non-Passers			Placement Passers			Credit Passers			Total		
	N	Rel. %	Cum. %	N	Rel. %	Cum. %	N	Rel. %	Cum. %	N	Rel. %	Cum.%
16 and Under	1	0.8	0.8	1	0.2	0.2	1	0.4	0.4	3	0.4	0.4
17	2	1.6	2.3	21	5.2	5.4	17	6.0	6.3	40	4.9	5.2
18	86	66.7	69.0	301	74.0	79.4	194	68.1	74.4	581	70.8	76.0
19	25	19.4	88.4	53	13.0	92.4	35	12.3	86.7	113	13.8	89.8
20	3	2.3	90.7	5	1.2	93.6	5	1.8	88.4	13	1.6	91.4
21	2	1.6	92.2	5	1.2	94.8	4	1.4	89.8	11	1.3	92.7
22	0	0.0	92.2	3	0.7	95.6	5	1.8	91.6	8	1.0	93.7
23	1	0.8	93.0	1	0.2	95.8	7	2.5	94.0	9	1.1	94.8
24	1	0.8	93.8	3	0.7	96.6	3	1.1	95.1	7	0.9	95.6
25-29	1	0.8	94.6	6	1.5	98.0	7	2.5	97.5	14	1.7	97.3
30-34	3	2.3	96.9	1	0.2	98.3	4	1.4	98.9	8	1.0	98.3
35 and Over	4	3.1	100.0	7	1.7	100.0	3	1.1	100.0	14	1.7	100.0
All Ages	129*	100.0	–	407	100.0	–	285*	100.0	–	821*	100.0	–

*Age and/or sex data unobtainable from 74 applicants

TCSUC –Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES

English Equivalency Examination – Spring 1977

Female Applicants by Age

Age	Non-Passers			Placement Passers			Credit Passers			Total		
	N	Rel. %	Cum. %	N	Rel. %	Cum. %	N	Rel. %	Cum.%	N	Rel. %	Cum. %
16 and Under	0	0.0	0.0	3	0.5	0.5	4	0.7	0.7	7	0.5	0.5
17	10	9.8	9.8	52	8.0	8.5	48	8.4	9.1	110	8.3	8.8
18	73	71.6	81.4	499	77.1	85.6	402	70.0	79.1	974	73.6	82.5
19	14	13.7	95.1	57	8.8	94.4	44	7.7	86.8	115	8.7	91.2
20	0	0.0	95.1	5	0.8	95.2	9	1.6	88.3	14	1.1	92.2
21	0	0.0	95.1	8	1.2	96.4	8	1.4	89.7	16	1.2	93.4
22	1	1.0	96.1	0	0.0	96.4	10	1.7	91.5	11	0.8	94.3
23	1	1.0	97.1	2	0.3	96.8	7	1.2	92.7	10	0.8	95.0
24	0	0.0	97.1	1	0.2	96.9	5	0.9	93.6	6	0.5	95.5
25-29	0	0.0	97.1	6	0.9	97.8	11	1.9	95.5	17	1.3	96.7
30-34	0	0.0	97.1	0	0.0	97.8	6	1.0	96.5	6	0.5	97.2
35 and Over	3	2.9	100.0	14	2.2	100.0	20	3.5	100.0	37	2.8	100.0
All Ages	102*	100.0	–	647*	100.0	–	574*	100.0	–	1,323*	100.0	–

*Age and/or sex data unobtainable from 74 applicants

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Converted Essay Scores
Of The Total Participant Population

Score	Frequency	Relative Percent	Cumulative Percent
22-24	1	0.0	0.0
25-29	31	1.4	1.4
30-34	245	11.0	12.5
35-39	266	12.0	24.5
40-44	717	32.3	56.8
45-49	308	13.9	70.7
50-54	238	10.7	81.4
55-59	303	13.7	95.1
60-64	62	2.8	97.9
65-69	36	1.6	99.5
70-74	9	0.4	99.9
75-79	2	0.0	100.0
Total	2,218	100.0	

Mean = 44.8 Median = 43 Standard Deviation = 8.8

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Objective Scores Of
The Total Participant Population

Score	Frequency	Relative Percent	Cumulative Percent
23-24	4	0.2	0.2
25-29	30	1.4	1.5
30-34	128	5.8	7.3
35-39	303	13.7	21.0
40-44	453	20.4	41.4
45-49	448	20.2	61.6
50-54	373	16.8	78.4
55-59	254	11.5	89.9
60-64	145	6.5	96.4
65-69	64	2.9	99.3
70-74	15	0.7	100.0
75-77	1	0.0	100.0
Total	2,218	100.0	

Mean = 47.1 Median = 47 Standard Deviation = 9.1

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Composite Scores Of The Total Participant Population

Score	Frequency	Relative Percent	Cumulative Percent
50- 59	26	1.2	1.2
60- 69	105	4.7	5.9
70- 79	370	16.7	22.6
80- 89	548	24.7	47.3
90- 99	502	22.6	69.9
100-109	361	16.3	86.2
110-119	189	8.5	94.7
120-129	89	4.0	98.7
130-139	25	1.1	99.9
140-149	3	0.1	100.0
Total	2,218	100.0	

Mean = 91.9 Median = 90 Standard Deviation = 15.6

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Converted Essay Scores
Of Those Participants Who Passed For Credit

Score	Frequency	Relative Percent	Cumulative Percent
43	131	14.8	14.8
47	201	22.7	37.6
51	187	21.2	58.7
55	170	19.2	77.9
59	90	10.2	88.1
63	61	6.9	95.0
66	24	2.7	97.7
69	10	1.1	98.9
72	8	0.9	99.8
76	2	0.2	100.0
Total	884	100.0	

Mean = 52.2 Median = 51 Standard Deviation = 6.8

TCSUC – Institutional Research
July 1977

78

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Objective Scores Of Those Participants Who Passed For Credit

Score	Frequency	Relative Percent	Cumulative Percent
45-47	119	13.5	13.5
48-50	171	19.3	32.8
51-53	153	17.3	50.1
54-56	113	12.8	62.9
57-59	126	14.3	77.1
60-62	82	9.3	86.4
63-65	72	8.1	94.6
66-68	28	3.2	97.7
69-71	15	1.7	99.4
72-74	4	0.5	99.9
75-77	1	0.1	100.0
Total	884	100.0	

Mean = 54.6 Median = 53 Standard Deviation = 6.3

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Composite Scores Of Those Participants Who Passed For Credit

Score	Frequency	Relative Percent	Cumulative Percent
90- 92	24	2.7	2.7
93- 96	115	13.0	15.7
97-100	157	17.8	33.5
101-104	138	15.6	49.1
105-109	146	16.5	65.6
110-114	111	12.6	78.2
115-119	76	8.6	86.8
120-124	63	7.1	93.9
125-129	26	2.9	96.8
130-134	17	1.9	98.8
135-139	8	0.9	99.7
140-144	3	0.3	100.0
Total	884	100.0	

Mean = 106.8 Median = 105 Standard Deviation = 10.3

TCSUC – Institutional Research
July 1977

80

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Converted Essay Scores
Of Participants Who Did Not Pass For Credit

Score	Frequency	Relative Percent	Cumulative Percent
22-24	1	0.1	0.1
25-29	31	2.3	2.4
30-34	245	18.4	20.8
35-39	266	19.9	40.7
40-44	586	43.9	84.6
45-49	107	8.0	92.7
50-54	51	3.8	96.5
55-59	43	3.2	99.7
60-64	1	0.1	99.8
65-69	2	0.1	99.9
70-72	1	0.1	100.0
Total	1,334	100.0	

Mean = 39.9 Median = 40 Standard Deviation = 6.1

81

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Objective Scores Of Participants Who Did Not Pass For Credit

Score	Frequency	Relative Percent	Cumulative Percent
23-24	4	0.3	0.3
25-29	30	2.2	2.5
30-34	128	9.6	12.1
35-39	303	22.7	34.9
40-44	453	34.0	68.6
45-49	238	17.8	86.7
50-54	108	8.1	94.8
55-59	47	3.5	98.3
60-64	16	1.2	99.5
65-69	6	0.4	99.9
70-73	1	0.1	100.0
Total	1,334	100.0	

Mean = 42.1 Median = 42 Standard Deviation = 7.0

82

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Composite Scores Of Participants Who Did Not Pass For Credit

Score	Frequency	Relative Percent	Cumulative Percent
50- 59	26	1.9	1.9
60- 69	105	7.9	9.8
70- 79	370	27.7	37.6
80- 89	548	41.1	78.6
90- 99	255	19.1	97.8
100-109	28	2.1	99.9
110-115	2	0.1	100.0
Total	1,334	100.0	

Mean = 82.0 Median = 83 Standard Deviation = 9.5

83

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Converted Essay Scores
Of Participants Who Passed For Placement

Score	Frequency	Relative Percent	Cumulative Percent
34	112	5.7	5.7
35-39	236	12.0	17.6
40-44	682	34.6	52.2
45-49	299	15.2	67.4
50-54	235	11.9	79.3
55-59	300	15.2	94.5
60-64	62	3.1	97.6
65-69	36	1.8	99.4
70-74	9	0.5	99.9
75-76	2	0.1	100.0
Total	1,973	100.0	

Mean = 46.1 Median = 43 Standard Deviation = 8.2

84

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Objective Scores Of
Participants Who Passed For Placement

Score	Frequency	Relative Percent	Cumulative Percent
35-39	268	13.6	13.6
40-44	431	21.8	35.4
45-49	434	22.0	57.4
50-54	367	18.6	76.0
55-59	250	12.7	88.7
60-64	144	7.3	96.0
65-69	63	3.2	99.2
70-74	15	0.8	99.9
75-77	1	0.1	100.0
Total	1,973	100.0	

Mean = 48.6 Median = 48 Standard Deviation = 8.2

TCSUC – Institutional Research
July 1977

85

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Composite Scores Of Participants Who Passed For Placement

Score	Frequency	Relative Percent	Cumulative Percent
69	8	0.4	0.4
70- 79	271	13.7	14.1
80- 89	527	26.7	40.9
90- 99	500	25.3	66.2
100-109	361	18.3	84.5
110-119	189	9.6	94.2
120-129	89	4.5	98.6
130-139	25	1.3	99.8
140-143	3	0.2	100.0
Total	1,973	100.0	

Mean = 94.7 Median = 93 Standard Deviation = 14.0

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Converted Essay Scores Of
Participants Who Did Not Pass For Credit Or Placement

Score	Frequency	Relative Percent	Cumulative Percent
22-24	1	0.4	0.4
25-29	31	12.7	13.1
30-34	133	54.3	67.3
35-39	30	12.2	79.6
40-44	35	14.3	93.9
45-49	9	3.7	97.6
50-54	3	1.2	98.8
55-59	3	1.2	100.0
Total	245	100.0	

Mean = 34.4 Median = 32 Standard Deviation = 5.8

TCSUC – Institutional Research
July 1977

87

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Objective Scores Of
Participants Who Did Not Pass For Credit Or Placement

Score	Frequency	Relative Percent	Cumulative Percent
23-24	4	1.6	1.6
25-29	30	12.2	13.9
30-34	128	52.2	66.1
35-39	35	14.3	80.4
40-44	22	9.0	89.4
45-49	14	5.7	95.1
50-54	6	2.4	97.6
55-59	4	1.6	99.2
60-64	1	0.4	99.6
65-66	1	0.4	100.0
Total	245	100.0	

Mean = 34.9 Median = 33.1 Standard Deviation = 6.9

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Distribution Of Composite Scores Of
Participants Who Did Not Pass For Credit Or Placement

Score	Frequency	Relative Percent	Cumulative Percent
50-54	6	2.4	2.4
55-59	20	8.2	10.6
60-64	41	16.7	27.3
65-69	56	22.9	50.2
70-74	66	26.9	77.1
75-79	33	13.5	90.6
80-84	14	5.7	96.3
85-89	7	2.9	99.2
90-94	1	0.4	99.6
95-96	1	0.4	100.0
Total	245	100.0	

Mean = 69.3 Median = 69 Standard Deviation = 8.1

TCSUC – Institutional Research
July 1977

83

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Intercorrelations Among Grouped Essay And Composite Scores
For The Total Applicant Population

	Question 1	Question 2	Essay Total	Conv. Essay	Objective	Composite
Question 1		0.4197	0.8157	0.8020	0.3941	0.6796
Question 2			0.8674	0.8641	0.4960	0.7739
Essay Total				0.9904	0.5321	0.8658
Converted Essay					0.5268	0.8681
Objective						0.8792

TCSUC – Institutional Research
July 1977

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
English Equivalency Examination – Spring 1977

Intercorrelations Among Part And Composite Scores
For The Total Applicant Population

	Essay #1		Essay #2					
	Read. 1	Read. 2	Read. 1	Read. 2	Essay Total	Essay Conv.	Objective	Composite
Essay #1								
Read. 1		0.5112	0.3173	0.3375	0.7078	0.6958	0.3385	0.5876
Read. 2			0.3221	0.3340	0.7103	0.6986	0.3465	0.5939
Essay #2								
Read. 1				0.6121	0.7765	0.7743	0.4316	0.6860
Read. 2					0.7812	0.7774	0.4596	0.7041
Essay Total						0.9904	0.5321	0.8658
Essay Conv.							0.5268	0.8681
Objective								0.8792

TCSUC – Institutional Research
July 1977